

Sous la direction de
Frédéric Bertrand et Myriam Maumy-Bertrand
Sandie Ferrigno
Didier Marx
Aurélie Muller-Gueudin

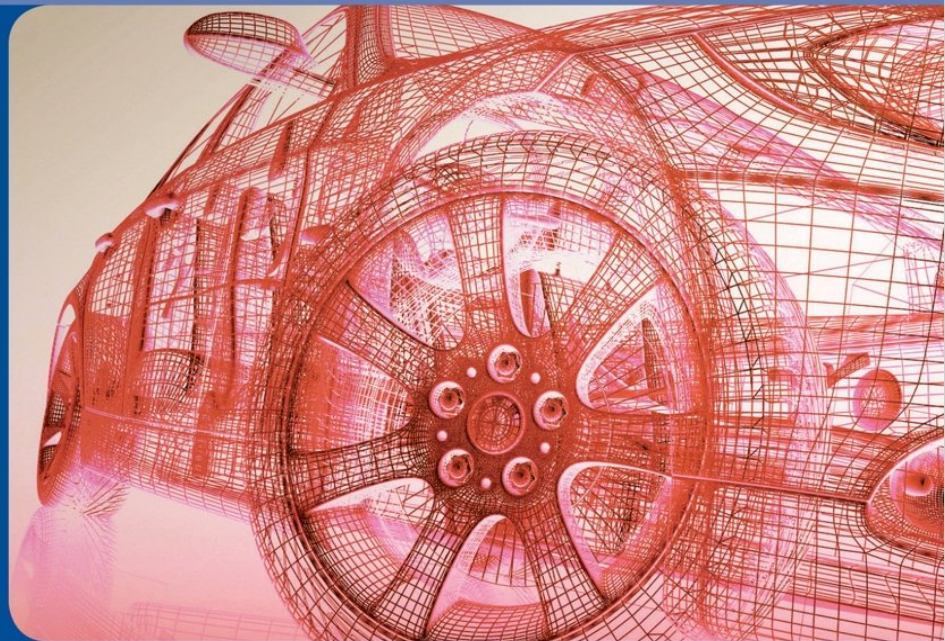
MATHÉMATIQUES POUR LES SCIENCES DE L'INGÉNIEUR

TOUT LE COURS EN FICHES

Licence • Prépas • IUT

- 120** Fiches de cours
- 120** exercices d'application
- 500** exercices d'entraînement

Avec ce livre,
des bonus sur
le web



DUNOD



SHOUT
SHARE **IT**

Tout le catalogue sur
www.dunod.com



Illustration de couverture :
3d model cars © 3ddock - Fotolia.com

Le pictogramme qui figure ci-contre mérite une explication. Son objet est d'alerter le lecteur sur la menace que représente pour l'avenir de l'écrit, particulièrement dans le domaine de l'édition technique et universitaire, le développement massif du photocopillage.

Le Code de la propriété intellectuelle du 1^{er} juillet 1992 interdit en effet expressément la photocopie à usage collectif sans autorisation des ayants droit. Or, cette pratique s'est généralisée dans les établissements



d'enseignement supérieur, provoquant une baisse brutale des achats de livres et de revues, au point que la possibilité même pour

les auteurs de créer des œuvres nouvelles et de les faire éditer correctement est aujourd'hui menacée. Nous rappelons donc que toute reproduction, partielle ou totale, de la présente publication est interdite sans autorisation de l'auteur, de son éditeur ou du

Centre français d'exploitation du droit de copie (CFC, 20, rue des Grands-Augustins, 75006 Paris).

© Dunod, Paris, 2013
ISBN 978-2-10-059755-0

Le Code de la propriété intellectuelle n'autorisant, aux termes de l'article L. 122-5, 2° et 3° a), d'une part, que les « copies ou reproductions strictement réservées à l'usage privé du copiste et non destinées à une utilisation collective » et, d'autre part, que les analyses et les courtes citations dans un but d'exemple et d'illustration, « toute représentation ou reproduction intégrale ou partielle faite sans le consentement de l'auteur ou de ses ayants droit ou ayants cause est illicite » (art. L. 122-4).

Cette représentation ou reproduction, par quelque procédé que ce soit, constituerait donc une contrefaçon sanctionnée par les articles L. 335-2 et suivants du Code de la propriété intellectuelle.

Table des matières

<i>Avant-propos</i>	IX
<i>Comment utiliser cet ouvrage ?</i>	X

Partie 1 Algèbre

Fiche 1	Logique	2
Fiche 2	Quantificateurs et raisonnements mathématiques	6
Fiche 3	Ensembles	10
Fiche 4	Relations binaires	14
Fiche 5	Applications	18
Fiche 6	Nombres entiers, nombres rationnels	22
Fiche 7	Structures algébriques : groupes	26
Fiche 8	Structures algébriques : anneaux et corps	30
Fiche 9	Arithmétique dans $\mathbb{Z}$	34
Fiche 10	Vecteurs	38
Fiche 11	Vecteurs et éléments de géométrie	42
Fiche 12	Polynômes	46
Fiche 13	Fractions rationnelles	50
Fiche 14	Systèmes linéaires	54
Fiche 15	Pivot de Gauss	58
Fiche 16	Nombres complexes	62
Fiche 17	Nombres complexes et géométrie plane	66
Fiche 18	Espaces vectoriels	70
Fiche 19	Bases – Dimension finie	74
Fiche 20	Applications linéaires	78
Fiche 21	Noyau et image d’une application linéaire	82
Fiche 22	Calcul matriciel	86
Fiche 23	Matrices et applications linéaires	90
Fiche 24	Déterminant	94
Fiche 25	Applications du calcul de déterminant	98
Fiche 26	Diagonalisation	102
Fiche 27	Applications de la diagonalisation	106
Fiche 28	Espaces préhilbertiens	110
Fiche 29	Orthogonalité, groupe orthogonal	114
Fiche 30	Coniques	118
<i>Focus</i>	<i>Matrices et cryptographie</i>	122

Partie 2

Analyse

Fiche 31	Nombres réels	126
Fiche 32	Suites numériques	130
Fiche 33	Convergence et divergence d'une suite numérique	134
Fiche 34	Suites arithmétiques et géométriques	138
Fiche 35	Suites particulières	142
Fiche 36	Continuité d'une fonction	146
Fiche 37	Dérivabilité d'une fonction	150
Fiche 38	Étude globale des fonctions dérivables	154
Fiche 39	Fonctions circulaires et circulaires réciproques	158
Fiche 40	Fonctions hyperboliques et hyperboliques réciproques	162
Fiche 41	Formules de Taylor	166
Fiche 42	Développements limités	170
Fiche 43	Étude d'une fonction	174
Fiche 44	Intégrales définies sur segment	178
Fiche 45	Primitives et intégrales d'une fonction continue	182
Fiche 46	Séries numériques	186
Fiche 47	Séries à termes positifs	190
Fiche 48	Suites de fonctions	194
Fiche 49	Séries de fonctions	198
Fiche 50	Séries entières	202
Fiche 51	Développement d'une fonction en série entière	206
Fiche 52	Séries de Fourier	210
Fiche 53	Intégration sur un intervalle quelconque	214
Fiche 54	Convergences monotone et dominée – Intégrales dépendant d'un paramètre	218
Fiche 55	Équations différentielles : premier ordre	222
Fiche 56	Équations différentielles : second ordre	226
Fiche 57	Fonction d'une variable réelle : Exercices de synthèse	230
Fiche 58	Fonctions de plusieurs variables	234
Fiche 59	Dérivées partielles – Gradient – Différentielle	238
Fiche 60	Dérivées partielles – Gradient – Différentielle (suite)	242
Fiche 61	Dérivées partielles d'ordre deux – Optimisation	246
Fiche 62	Courbes et surfaces paramétrées	250
Fiche 63	Courbes planes paramétrées	254
Fiche 64	Courbe plane définie par son équation polaire	258
Fiche 65	Intégrales doubles et multiples	262
Fiche 66	Sommabilité et intégrales doubles ou multiples généralisées	266
Fiche 67	Intégrales curvilignes	270
Fiche 68	Intégrales de surface	274
Fiche 69	Transformée de Laplace	278

Fiche 70	Exemples d'équations aux dérivées partielles	282
<i>Focus</i>	<i>Le Wi-Fi</i>	286

Partie 3 Probabilités

Fiche 71	Dénombrement	290
Fiche 72	Événements et probabilité	294
Fiche 73	Probabilité sur un univers fini, dénombrable ou continu	298
Fiche 74	Événements indépendants, Probabilité conditionnelle Théorème de Bayes	302
Fiche 75	Variable aléatoire réelle – Loi d'une variable aléatoire réelle	306
Fiche 76	Fonction de répartition	310
Fiche 77	Espérance	314
Fiche 78	Moments, Variance et Écart-type	318
Fiche 79	Paramètres de position – Intervalles de probabilité	322
Fiche 80	Fonction d'une variable aléatoire réelle – Calcul de loi	326
Fiche 81	Lois discrètes usuelles à valeurs dans un ensemble fini	330
Fiche 82	Lois discrètes usuelles à valeurs dans un ensemble dénombrable	334
Fiche 83	Loi normale	338
Fiche 84	Lois continues usuelles	342
Fiche 85	Lois continues usuelles (suite)	346
Fiche 86	Couple aléatoire – Loi d'un couple discret	350
Fiche 87	Loi d'un couple aléatoire continu	354
Fiche 88	Fonction de répartition d'un couple aléatoire	358
Fiche 89	Moments de plusieurs variables aléatoires réelles	362
Fiche 90	Variables aléatoires indépendantes	366
Fiche 91	Fonction d'un couple aléatoire – Calcul de loi	370
Fiche 92	Loi des grands nombres	374
Fiche 93	Théorème de la limite centrale	378
Fiche 94	Approximations d'une loi par une autre	382
Fiche 95	Formulaire	386
<i>Focus</i>	<i>Probabilités et jeux de hasard</i>	388

Partie 4 Statistique

Fiche 96	Concepts fondamentaux de la statistique	392
Fiche 97	Statistique descriptive univariée	396
Fiche 98	Représentations graphiques pour les séries statistiques quantitatives	400
Fiche 99	Représentations graphiques pour les séries statistiques qualitatives	404

Fiche 100	Caractéristiques de position	408
Fiche 101	Caractéristiques de dispersion	412
Fiche 102	Boîte à moustaches et caractéristiques de forme	416
Fiche 103	Statistique descriptive bivariable	420
Fiche 104	Représentations graphiques bivariées	424
Fiche 105	Mesures de liaison entre deux caractères	428
Fiche 106	Échantillonnage – Modèles : vocabulaire	432
Fiche 107	Estimateurs et propriétés	434
Fiche 108	Méthodes de construction d'estimateurs	438
Fiche 109	Exemples d'estimateurs de caractéristiques de position	442
Fiche 110	Exemples d'estimateurs de caractéristiques de dispersion	446
Fiche 111	Estimation par intervalle de confiance et intervalle de confiance pour une proportion	450
Fiche 112	Intervalles de confiance pour une espérance et une variance	454
Fiche 113	Introduction à la théorie des tests d'hypothèses	458
Fiche 114	Tests de conformité d'une espérance ou d'une variance à une norme	462
Fiche 115	Tests de comparaison de deux variances	466
Fiche 116	Tests de comparaison de deux espérances	470
Fiche 117	Tests du Khi-deux : adéquation et indépendance	474
Fiche 118	Test de normalité de Shapiro-Wilk. Diagramme quantile-quantile	478
Fiche 119	Régression linéaire simple	482
Fiche 120	Intervalles de confiance et tests en régression linéaire simple	486
<i>Annexes</i>		490
<i>Focus</i>	<i>Les sondages</i>	502

Corrigés des exercices

505

(Les corrigés d'une sélection d'exercices sont disponibles sur dunod.com)

<i>Index</i>	557
--------------	-----

Avant-propos

Cet ouvrage résulte de la collaboration de quatre mathématiciens (Aurélie, Frédéric, Myriam et Sandie) et d'un physicien (Didier). Il est organisé en quatre parties : algèbre, analyse, probabilités et statistique. Il s'adresse aux étudiants des deux premières années post-bac : licences, prépas intégrées ou IUT. Ce livre peut également aider à la préparation au CAPES de mathématiques.

Le cours, structuré en fiches, est exposé de manière claire et synthétique. Chaque fiche présente les points essentiels à retenir, des exercices d'application illustrent les notions utiles et de nombreux exercices corrigés permettent de se préparer aux examens et concours. Certains corrigés sont disponibles sur le site dunod.com sur la page de description de l'ouvrage. Quatre focus apportent enfin des compléments historiques ou techniques en lien avec des sujets de société.

Dans la collection « Tout le cours en fiches », vous trouverez donc l'essentiel, sauf votre propre travail, bien sûr. Alors courage !

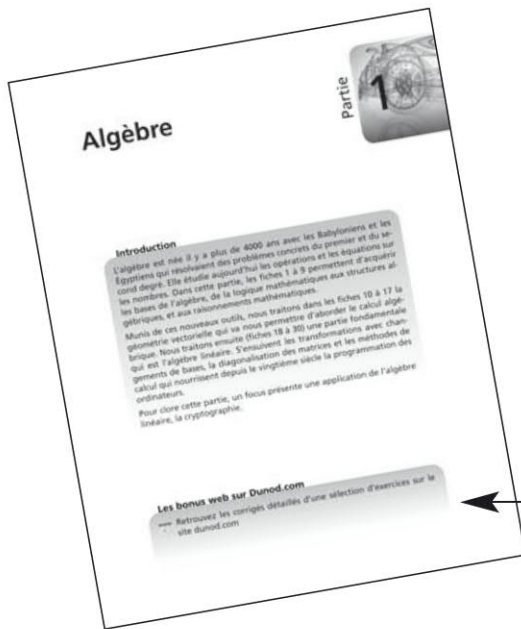
Toutes vos remarques, vos commentaires, vos critiques, et même vos encouragements seront accueillis avec plaisir aux adresses électroniques suivantes :

fbertran@math.unistra.fr
mmaumy@math.unistra.fr

Ce livre a bénéficié de la relecture attentive d'un comité de relecteurs, constitué d'enseignants et d'étudiants. Nous souhaitons qu'ils soient tous grandement remerciés ici :

Samuela Aubin, maître de conférences à l'INSA de Lyon ;
Jean-Romain Heu, professeur agrégé à l'INSA de Strasbourg ;
Jean Labourdette, docteur en mathématique et directeur de l'ESIEA Ouest ;
Vincent Lécuyer, professeur agrégé à l'ENSIC de Nancy ;
James Ledoux, professeur à l'INSA de Rennes ;
Renaud Marty et Bruno Pinçon, tous deux maîtres de conférences à l'université de Lorraine ;
Constantin Morarescu, maître de conférences à l'ENSEM de Nancy ;
Valentin Savidan, étudiant en DUAS 1 à l'université de Strasbourg ;
Raphaële Supper, maître de conférences à l'université de Strasbourg.

Comment utiliser cet ouvrage ?



Un découpage
en quatre grandes parties :
Algèbre, Analyse, Probabilités,
Statistique

Des compléments sur dunod.com

120 fiches de cours
Les notions essentielles avec des renvois
pour naviguer d'une fiche à l'autre

De très
nombreux
exemples

Un repérage
facile

fiche 1 Logique

La logique mathématique permet l'étude des mathématiques en tant que langage.

Définition 1.1

Une **assertion** est un énoncé mathématique auquel nous pouvons attribuer la valeur de vérité vrai (V) ou faux (F) mais jamais les deux simultanément.

Exemples

1. L'assertion « $10 > 100$ » est fautive (F).
2. L'assertion « 12 est un multiple de 4 » est vraie (V).
3. L'énoncé « π vaut approximativement $3,14$ » n'est pas une assertion car il n'est pas assez précis pour déterminer s'il est vrai ou faux. En effet, il sera vrai si nous nous contentons d'une approximation de la valeur de π au centième près et faux sinon.

Définition 1.2

Un **prédicat** P est un énoncé mathématique contenant des lettres appelées variables tel que, quand nous remplaçons chacune de ces variables par un élément donné de cette variable nous obtenons une assertion.

Exemple

L'énoncé suivant « x est un multiple de 6 » est un prédicat car il devient une assertion lorsque nous donnons une valeur à x . En effet,

- « 10 est un multiple de 6 » est une assertion fautive (F).
- « 12 est un multiple de 6 » est une assertion vraie (V).

Les **connecteurs logiques** permettent de créer de nouveaux prédicats, dits **composés**, à partir de prédicats de référence.

Définition 1.3

Soit P un prédicat. La **négation** de P est le prédicat noté $\text{non}(P)$ ou $\neg P$, qui est vrai lorsque P est faux et faux lorsque P est vrai. Nous résumons ceci dans la table de vérité suivante :

P	$\text{non}(P)$
V	F
F	V

Exemple

Soit P le prédicat « $x > 5$ ». Alors $\text{non}(P)$ est le prédicat « $x \leq 5$ ».

Remarque : En effet, le contraire de « supérieur à » est « inférieur ou égal à », et non simplement « inférieur à ». De même, le contraire de « pour tout x , P » n'est pas « quel que soit x , $\text{non}(P)$ » mais « il existe x pour lequel $\text{non}(P)$ ».

Définition 1.4

Soient P et Q deux prédicats.

- Le prédicat « $P \Rightarrow Q$ » appelé **implication** de P vers Q est un prédicat qui est faux lorsque P est vrai et Q est faux, et vrai dans tous les autres cas.
- Le prédicat « $P \Leftrightarrow Q$ » appelé **équivalence** de P et de Q est un prédicat qui est vrai lorsque P et Q sont simultanément vrais ou faux, et faux dans tous les autres cas.
- Le prédicat « $P \wedge Q$ » (P et Q) appelé **conjonction** de P et de Q est un prédicat qui est vrai lorsque P et Q sont simultanément vrais, et faux dans tous les autres cas.
- Le prédicat « $P \vee Q$ » (P ou Q) appelé **disjonction** de P et de Q est un prédicat qui est vrai lorsque au moins un des prédicats P et Q est vrai, et faux dans tous les autres cas.

Nous résumons ceci dans la table de vérité :

P	Q	$P \Rightarrow Q$	$P \Leftrightarrow Q$	$P \wedge Q$	$P \vee Q$
V	V	V	V	V	V
V	F	F	F	F	V
F	V	V	F	F	V
F	F	V	V	F	F

Le « ou » dans la définition de $P \vee Q$ a un sens inclusif, il ne se confond pas avec le « ou » exclusif de « fromage ou dessert ».

Théorème 1.1

- Soient P et Q deux prédicats. Nous avons les équivalences logiques (notées par $=$) suivantes :

$$\text{non}(P \text{ ou } Q) = (\text{non}(P) \text{ et } \text{non}(Q)),$$
$$\text{non}(P \text{ et } Q) = (\text{non}(P) \text{ ou } \text{non}(Q)).$$

Ce sont les **lois de Morgan** pour les prédicats.

- Soient P , Q et R trois prédicats. Nous avons les équivalences logiques suivantes :

$$(P \text{ ou } (Q \text{ et } R)) = ((P \text{ ou } Q) \text{ et } (P \text{ ou } R)),$$
$$(P \text{ et } (Q \text{ ou } R)) = ((P \text{ et } Q) \text{ ou } (P \text{ et } R)).$$

fiche 1

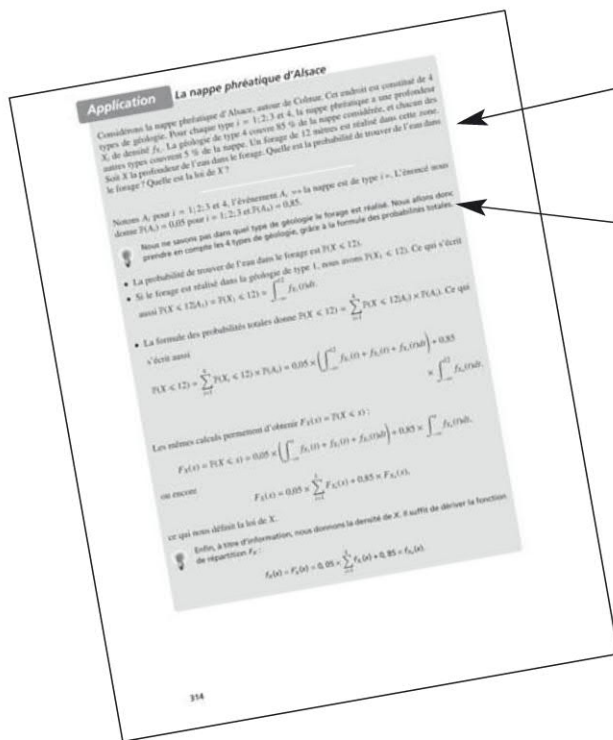
Algèbre

Analyse

Probabilités

Statistique

Solutions

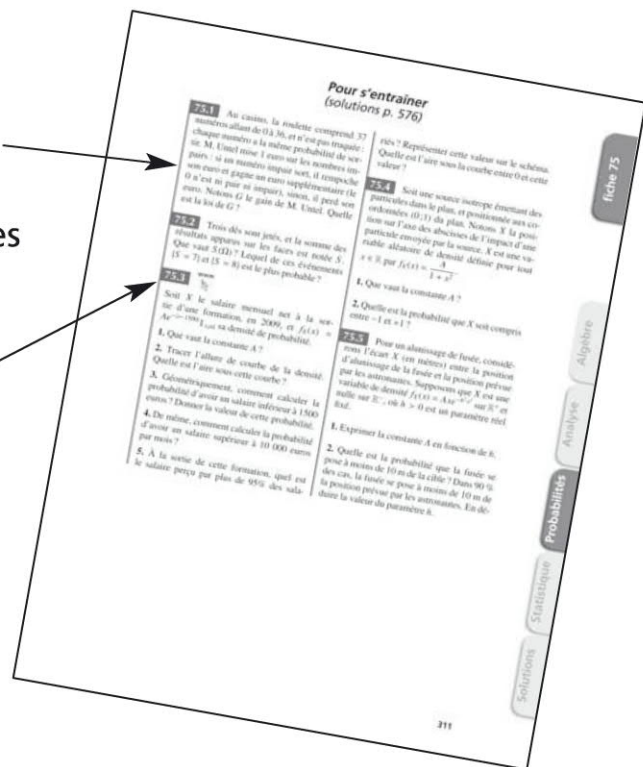


Des exemples d'applications dans tous les domaines des sciences

Des conseils méthodologiques

Des exercices pour s'entraîner
Les solutions sont regroupées
en fin d'ouvrage ou disponibles
sur le site dunod.com

Des renvois aux bonus web



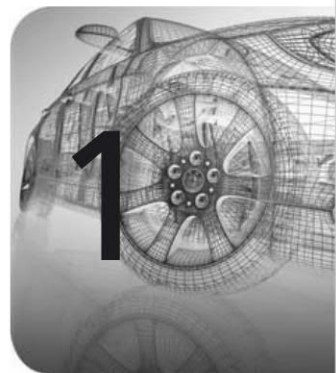
Retrouvez les bonus web sur dunod.com :

- Les corrigés détaillés d'une sélection d'exercices signalés par le pictogramme  www.dunod.com.
- Des informations sur le logiciel libre R.

Algèbre

Partie

1



Introduction

L'algèbre est née il y a plus de 4000 ans avec les Babyloniens et les Égyptiens qui résolvait des problèmes concrets du premier et du second degrés. Elle étudie aujourd'hui les opérations et les équations sur les nombres.

Dans cette partie, les fiches 1 à 9 permettent d'acquérir les bases de l'algèbre, de la logique mathématique aux structures algébriques.

Munis de ces nouveaux outils, nous traitons dans les fiches 10 à 17 la géométrie vectorielle qui va nous permettre d'aborder le calcul algébrique. Nous explorons ensuite (fiches 18 à 30) une partie fondamentale qui est l'algèbre linéaire. S'ensuivent les transformations avec changements de bases, la diagonalisation des matrices et les méthodes de calcul qui nourrissent depuis le xx^e siècle la programmation des ordinateurs.

Pour clore cette partie, un focus présente une application de l'algèbre linéaire : la cryptographie.

Les bonus web sur Dunod.com

 Retrouvez les corrigés détaillés d'une sélection d'exercices sur le site dunod.com

La logique mathématique permet l'étude des mathématiques en tant que langage.

Définition 1.1

Une **assertion** est un énoncé mathématique auquel nous pouvons attribuer la valeur de vérité vrai (V) ou faux (F) mais jamais les deux simultanément.

Exemples

1. L'assertion « $10 > 100$ » est fausse (F).
2. L'assertion « 12 est un multiple de 4 » est vraie (V).
3. L'énoncé « π vaut approximativement 3, 14 » n'est pas une assertion car il n'est pas assez précis pour déterminer s'il est vrai ou faux. En effet, il sera vrai si nous nous contentons d'une approximation de la valeur de π au centième près et faux sinon.

Définition 1.2

Un **prédicat** P est un énoncé mathématique contenant des lettres appelées variables tel que, quand nous remplaçons chacune de ces variables par un élément donné de cette variable nous obtenons une assertion.

Exemple

L'énoncé suivant « n est un multiple de 6 » est un prédicat car il devient une assertion lorsque nous donnons une valeur à n . En effet,

- « 10 est un multiple de 6 » est une assertion fausse (F).
- « 12 est un multiple de 6 » est une assertion vraie (V).



Les **connecteurs logiques** permettent de créer de nouveaux prédicats, dits **composés**, à partir de prédicats de référence.

Définition 1.3

Soit P un prédicat. La **négation** de P est le prédicat noté $\text{non}(P)$ ou $\neg P$, qui est vrai lorsque P est faux et faux lorsque P est vrai. Nous résumons ceci dans la table de vérité suivante :

P	$\text{non}(P)$
V	F
F	V

Exemple

Soit P le prédicat « $x > 5$ ». Alors $\text{non}(P)$ est le prédicat « $x \leq 5$ ».

Remarque : En effet, le contraire de « supérieur à » est « inférieur ou égal à », et non simplement « inférieur à ». De même, le contraire de « pour tout x , P » n'est pas « quel que soit x , $\text{non}(P)$ » mais « il existe x pour lequel $\text{non}(P)$ ».

Définition 1.4

Soient P et Q deux prédicats.

- Le prédicat « $P \Rightarrow Q$ » appelé **implication** de P vers Q est un prédicat qui est faux lorsque P est vrai et Q est faux, et vrai dans tous les autres cas.
- Le prédicat « $P \Leftrightarrow Q$ » appelé **équivalence** de P et de Q est un prédicat qui est vrai lorsque P et Q sont simultanément vrais ou faux, et faux dans tous les autres cas.
- Le prédicat « $P \wedge Q$ » (P et Q) appelé **conjonction** de P et de Q est un prédicat qui est vrai lorsque P et Q sont simultanément vrais, et faux dans tous les autres cas.
- Le prédicat « $P \vee Q$ » (P ou Q) appelé **disjonction** de P et de Q est un prédicat qui est vrai lorsque au moins un des prédicats P et Q est vrai, et faux dans tous les autres cas.

Nous résumons ceci dans la table de vérité :

P	Q	$P \Rightarrow Q$	$P \Leftrightarrow Q$	$P \wedge Q$	$P \vee Q$
V	V	V	V	V	V
V	F	F	F	F	V
F	V	V	F	F	V
F	F	V	V	F	F



Le « ou » dans la définition de $P \vee Q$ a un sens inclusif, à ne pas confondre avec le « ou » exclusif de « fromage ou dessert ».

Théorème 1.1

- Soient P et Q deux prédicats. Nous avons les équivalences logiques (notées par $=$) suivantes :

$$\text{non}(P \text{ ou } Q) = (\text{non}(P) \text{ et } \text{non}(Q)),$$

$$\text{non}(P \text{ et } Q) = (\text{non}(P) \text{ ou } \text{non}(Q)).$$

Ce sont les **lois de Morgan** pour les prédicats.

- Soient P , Q et R trois prédicats. Nous avons les équivalences logiques suivantes :

$$(P \text{ ou } (Q \text{ et } R)) = ((P \text{ ou } Q) \text{ et } (P \text{ ou } R)),$$

$$(P \text{ et } (Q \text{ ou } R)) = ((P \text{ et } Q) \text{ ou } (P \text{ et } R)).$$

Théorème 1.2

Soient P et Q deux prédicats. Nous avons les équivalences logiques suivantes :

- $\text{non}(\text{non}(P)) = P$,
- $[P \Rightarrow Q] = [\text{non}(P) \text{ ou } Q]$,

- $[\text{non}(P \Rightarrow Q)] = [P \text{ et non}(Q)]$,
- $[P \Rightarrow Q] = [\text{non}(Q) \Rightarrow \text{non}(P)]$,
- $[P \Leftrightarrow Q] = [(P \Rightarrow Q) \text{ et } (Q \Rightarrow P)]$.



$P \Rightarrow Q$ signifie que Q est une condition nécessaire pour P . L'implication $\text{non}(Q) \Rightarrow \text{non}(P)$ est la **contraposée** de $P \Rightarrow Q$. L'implication $Q \Rightarrow P$ est la **réciproque** de $P \Rightarrow Q$. Pour démontrer une équivalence, nous démontrons souvent une implication et sa réciproque.

Définition 1.5

Une **tautologie** est une formule propositionnelle qui prend la valeur « vrai » quelles que soient les valeurs données à ses arguments.

Exemples

Les formules suivantes sont des tautologies :

1. $P \text{ ou non}(P)$, qui est le **Principe du tiers exclu**,
2. $P \Rightarrow P$,
3. $(P \text{ ou } Q) \Rightarrow P$
4. $P \Rightarrow (P \text{ et } Q)$

Définition 1.6

Une **contradiction** est une formule qui prend la valeur « faux » quelles que soient les valeurs données à ses arguments.

Exemple

$P \text{ et non}(P)$ est une contradiction. Sa négation, qui est donc une tautologie, est le **principe de non-contradiction**.

Application

Nous suspectons trois élèves, Anita, Bernard et Christophe d'avoir dérobé du matériel dans un laboratoire de Physique. Nous possédons les informations suivantes à leur sujet :

- Si Christophe n'est pas coupable alors Bernard est coupable,
- Si Anita n'est pas coupable alors Christophe est coupable,
- Si Christophe est coupable alors Anita l'est aussi,
- Si Anita est coupable alors Bernard ne l'est pas.

Les assertions suivantes sont-elles vraies ou fausses ?

- C : « Christophe est coupable »,
- B : « Bernard est coupable »,
- A : « Anita est coupable »,
- D : « Anita ou Bernard est coupable ».



- La première information de l'énoncé se traduit par $\text{non}(C) \Rightarrow B$. Donc $\text{non}(B) \Rightarrow C$ par contraposée.
- La deuxième information se traduit par $\text{non}(A) \Rightarrow C$. Donc $\text{non}(C) \Rightarrow A$ par contraposée.
- La troisième information se traduit par $C \Rightarrow A$. Donc $\text{non}(A) \Rightarrow \text{non}(C)$ par contraposée.
- Enfin, la dernière information se traduit par $A \Rightarrow \text{non}(B)$. Donc $B \Rightarrow \text{non}(A)$ par contraposée.

Nous en déduisons que :

- L'assertion C : « Christophe est coupable » est vraie.
En effet, si nous supposons que $\text{non}(C)$ est vrai, $\text{non}(C) \Rightarrow B$, alors $B \Rightarrow \text{non}(A)$ et $\text{non}(A) \Rightarrow C$. Nous avons alors $\text{non}(C) \Rightarrow C$ ce qui est contradictoire.
- L'assertion B : « Bernard est coupable » est fausse.
En effet, si nous supposons que B est vrai, $B \Rightarrow \text{non}(A)$, alors $\text{non}(A) \Rightarrow \text{non}(B)$. Nous aurions alors $B \Rightarrow \text{non}(B)$ ce qui est contradictoire.
- L'assertion A : « Anita est coupable » est vraie.
En effet, si nous supposons que $\text{non}(A)$ est vrai, $\text{non}(A) \Rightarrow C$, alors $C \Rightarrow A$ et $\text{non}(A) \Rightarrow A$ ce qui est impossible.
- Enfin, comme Bernard n'est pas coupable, l'assertion D : « Anita ou Bernard est coupable » ne sera vraie que si Anita est coupable, ce qui est le cas puisque nous avons montré que l'assertion A est vraie dans l'item précédent. Donc l'assertion D est vraie.

Pour s'entraîner (solutions p. 505)

1.1 Soit l'assertion « $x = 2$ ». Donner la négation de cette assertion.

1.2 Écrire les contraposées des implications suivantes :

1. $xy \neq 0 \Rightarrow x \neq 0$ et $y \neq 0$,
2. $[n \text{ premier}] \Rightarrow [n = 2 \text{ ou } n \text{ impair}]$.

1.3 Montrer que $[n \text{ pair}] \Leftrightarrow [n^2 \text{ pair}]$.

1.4 Nous considérons les assertions P et Q listées ci-dessous. Donner dans chaque cas la valeur de vérité pour $P \wedge Q$.

- P : « Paris est la capitale de l'Espagne », Q : « $2 + 2 = 4$ »,
- P : « Un chat ronronne », Q : « Un carré a quatre cotés égaux »,

- P : « $4 \times 6 = 21$ », Q : « Venise se situe en Italie »,

- P : « Deux droites parallèles se coupent en un point », Q : « Paris est une ville de moins de dix mille habitants ».

1.5 Nous considérons les assertions P et Q listées ci-dessous. Donner dans chaque cas la valeur de vérité pour $P \vee Q$.

- P : « Un oiseau sait nager », Q : « Paris est la capitale de la France »,
- P : « Un chien a cinq pattes », Q : « Un triangle a trois côtés »,
- P : « $2 < 3$ », Q : « Madrid est la capitale de l'Espagne »,
- P : « Une voiture a deux roues », Q : « $1 + 2 = 4$ ».

Quantificateurs et raisonnements mathématiques

1. Quantificateurs

À partir d'un prédicat P , défini sur un ensemble E , nous pouvons construire de nouvelles assertions, dites quantifiées, en utilisant les **quantificateurs** « quel que soit » et « il existe ».

Définition 2.1

Le quantificateur universel « **quel que soit** » ou « **pour tout** », noté par $\forall$, permet de définir l'assertion quantifiée « $\forall x \in E, P(x)$ » qui est vraie si pour tous les éléments $x \in E$, l'assertion $P(x)$ est vraie.

Exemple

$\forall n \in \mathbb{N}^*$, l'assertion $P : \langle n^2 > 0 \rangle$ est vraie.

Définition 2.2

Le quantificateur existentiel « **il existe** », noté par $\exists$, permet de définir l'assertion quantifiée « $\exists x \in E, P(x)$ » qui est vraie si nous pouvons trouver au moins un élément $x \in E$ tel que l'assertion $P(x)$ soit vraie. S'il en existe un et un seul nous pourrions écrire « $\exists! x \in E, P(x)$ » et nous dirions qu'il existe un unique élément x de E vérifiant $P(x)$.

Exemple

L'assertion quantifiée $P : \langle \exists x \in \mathbb{R}, x^2 = 4 \rangle$ est vraie.



Si « $\forall x \in E, P(x)$ » est vraie alors « $\exists x \in E, P(x)$ » est vraie.

Définition 2.3

Soit $P(x)$ un prédicat sur un ensemble E . Alors,

$$[\text{non}(\forall x \in E, P(x))] = [\exists x \in E, \text{non}(P(x))],$$

$$[\text{non}(\exists x \in E, P(x))] = [\forall x \in E, \text{non}(P(x))].$$



Soient E et F deux ensembles. Soit $P(x, y)$ un prédicat à deux variables avec $x \in E$ et $y \in F$.

- L'assertion quantifiée « $\forall x \in E, \forall y \in F, P(x, y)$ » est vraie lorsque tous les éléments x de E et tous les éléments y de F vérifient $P(x, y)$.
- L'assertion quantifiée « $\exists x \in E, \exists y \in F, P(x, y)$ » est vraie lorsqu'il existe au moins un élément x appartenant à E et au moins un élément y appartenant à F vérifiant $P(x, y)$.



Si nous utilisons deux fois le même quantificateur, l'ordre n'a pas d'importance, nous pouvons alors les permuter. En revanche, si les quantificateurs sont différents, leur ordre est important.

2. Raisonnements mathématiques

Nous donnons dans ce paragraphe quelques méthodes de démonstrations basées sur des raisonnements mathématiques que nous serons amenés à utiliser dans les prochaines fiches de cet ouvrage.

► Raisonnement par hypothèse auxiliaire ou déduction

Le but est de démontrer qu'un énoncé Q est vrai. Si l'énoncé P est vrai et si l'implication $P \Rightarrow Q$ est vraie alors l'énoncé Q est nécessairement vrai.

► Raisonnement par l'absurde

Le but est de démontrer qu'un énoncé P est vrai. Un raisonnement par l'absurde consiste à montrer que $\text{non}(P)$ entraîne à la fois un énoncé Q et son contraire $\text{non}(Q)$. Nous supposons l'énoncé $\text{non}(P)$ vrai et nous cherchons alors Q qui, sous cette hypothèse, serait à la fois vrai et faux. Nous disons que nous avons obtenu une contradiction ou que l'hypothèse $\text{non}(P)$ est contradictoire.

► Raisonnement par contraposée

Le but est de démontrer des résultats en faisant apparaître une implication « $P \Rightarrow Q$ ». Le principe est qu'au lieu de montrer « $P \Rightarrow Q$ » nous montrons sa contraposée « $\text{non}(Q) \Rightarrow \text{non}(P)$ ». Nous faisons l'hypothèse que $\text{non}(Q)$ est vraie et nous montrons que cela entraîne que $\text{non}(P)$ est vraie.

► Raisonnement par contre-exemple

Le raisonnement par contre-exemple sert à montrer qu'un énoncé de la forme « $\forall x \in E, P(x)$ » est un énoncé faux. Nous cherchons alors à trouver un élément x de E qui ne vérifie pas $P(x)$.

► Raisonnement par récurrence simple

Le raisonnement par récurrence simple sert à montrer qu'un énoncé de la forme « $\forall n \geq n_0, P(n)$ » est vrai. Si la propriété $P(n_0)$ est vraie et si l'implication « $P(n) \Rightarrow P(n+1)$ » est vraie pour tout $n \geq n_0$, alors la propriété $P(n)$ est vraie pour tout $n \geq n_0$. Il s'effectue en trois étapes :

- **Étape d'initialisation.** Nous vérifions que $P(n_0)$ est vraie.
- **Étape d'hérédité.** Fixons un entier naturel $n \geq n_0$, puis montrons que si $P(n)$ est vraie alors $P(n+1)$ est vraie.
- **Étape de conclusion.** Nous concluons que l'assertion $P(n)$ est vraie $\forall n \geq n_0$.



Il existe également le raisonnement par récurrence à deux pas et le raisonnement par récurrence forte. Pour le premier, nous supposons $P(n)$ et $P(n+1)$ vraies et nous montrons que $P(n+2)$ est vraie. Pour le second, nous supposons que pour tout $k \leq n$, $P(k)$ est vraie et nous montrons que $P(n+1)$ est vraie.

Application

1. Démontrer par l'absurde l'énoncé suivant : $\sqrt{2}$ est irrationnel.
2. Soient x et y deux réels. Montrer que

$$xy \neq 0 \implies x \neq 0 \text{ et } y \neq 0.$$

3. Montrer que

$$\forall a \in \mathbb{R}_+, \forall n \in \mathbb{N}, (1+a)^n \geq 1+na.$$

1. Si $\sqrt{2}$ est rationnel, nous pouvons écrire $\sqrt{2}$ sous la forme p/q , avec $p, q \in \mathbb{Z}$ et p et q premiers entre eux. Nous avons alors $p^2 = 2q^2$. Donc p^2 est pair ce qui implique que p l'est également. Donc si p est pair, alors p peut s'écrire sous la forme $2n$. Donc $q^2 = 2n^2$ et par conséquent q est aussi pair. Mais alors, p et q ne peuvent pas être premiers entre eux, ce qui contredit l'hypothèse. Ainsi, nous ne pouvons pas écrire $\sqrt{2}$ sous la forme p/q et donc $\sqrt{2}$ est bien irrationnel.

2.



Nous montrons ce résultat en utilisant un raisonnement par contraposée.

La contraposée de $xy \neq 0 \implies x \neq 0$ et $y \neq 0$ est

$$\text{si } x = 0 \text{ ou } y = 0, \text{ alors } xy = 0.$$

Montrons donc cette assertion. Si nous choisissons $x = 0$ ou $y = 0$, alors le produit xy est nécessairement nul. Nous en déduisons donc le résultat cherché.

3. Pour montrer ce résultat nous allons utiliser le raisonnement par récurrence. Nous allons procéder en trois étapes.

- **Étape d'initialisation**

Pour $n = 0$, nous avons $(1+a)^0 = 1$ qui est bien supérieur ou égal à $1+na = 1+0 \times a$.

- **Étape d'hérédité**

Nous supposons que la propriété est vraie à un ordre $n \in \mathbb{N}$, c'est-à-dire que $(1+a)^n \geq 1+na$. Nous voulons montrer que $(1+a)^{n+1} \geq 1+(n+1)a$. Nous avons, grâce à l'hypothèse d'hérédité,

$$\begin{aligned}(1+a)^{n+1} &= (1+a)(1+a)^n, \\ &\geq (1+a)(1+na), \\ &= 1+(n+1)a+na^2, \\ &\geq 1+(n+1)a,\end{aligned}$$

car $1+a > 0$ et $na^2 \geq 0$. Donc l'assertion est vraie à l'ordre $n+1$.

- **Étape de conclusion**

Nous en déduisons que l'assertion est vraie pour tout $n \in \mathbb{N}$.

Pour s'entraîner (solutions p. 505)

2.1 Montrer en utilisant un raisonnement par contraposée que pour n entier, si n^2 est impair alors n est impair.

2.2 Montrer en utilisant un raisonnement par contraposée que pour x et y réels différents de 1, si $x \neq y$ alors $\frac{1}{x-1} \neq \frac{1}{y-1}$.

2.3 Montrer, en donnant un contre-exemple, que $2^{2^n} + 1$ n'est pas un nombre premier pour toute valeur de $n \in \mathbb{N}$.

2.4  Montrer que la somme d'un nombre rationnel et d'un nombre irrationnel est irrationnelle.

2.5 Montrer par récurrence que, pour tout $n \in \mathbb{N}$, $2^n \geq n + 1$.

2.6 Montrer par récurrence que, pour tout $n \geq 4$, nous avons $2^n \leq n!$.

2.7 Montrer par récurrence que, pour tout $n \geq 1$, $3^{2n+2} - 2^{n+1}$ est divisible par 7.

2.8 Montrer par récurrence que, pour tout $n \geq 1$, $n(2n+1)(7n+1)$ est divisible par 6.

2.9 Montrer par récurrence que, pour tout $n \in \mathbb{N}^*$, l'égalité $\sum_{k=1}^n k = \frac{n(n+1)}{2}$ est vraie.

2.10 Montrer par récurrence que, pour tout $n \geq 1$, l'égalité $\sum_{k=1}^n (2k-1) = n^2$ est vraie.

2.11  Soient $a, b \geq 0$. Montrer que si $\frac{a}{1+b} = \frac{b}{1+a}$ alors $a = b$.

2.12  Montrer que l'assertion suivante est fausse : « Tout entier positif est somme de trois carrés. »
(Les carrés sont les $0^2; 1^2; 2^2; \dots$. Par exemple $b = 1^2 + 1^2 + 2^2$.)

1. Définitions

Définition 3.1

Un **ensemble** E est une collection d'objets telle que pour tout élément x , nous pouvons dire si l'assertion x appartient à E ($x \in E$) est vraie ou fausse. L'**ensemble vide**, c'est-à-dire l'ensemble qui ne contient aucun élément, est noté par $\emptyset$. L'ensemble ne contenant qu'un seul élément x est un **singleton**, noté $\{x\}$.

Exemples

1. $\mathbb{R}$ est l'ensemble des nombres réels.
2. $\mathbb{N}$ est l'ensemble des nombres entiers naturels.
3. $\{1; 2; 3; 4\}$ est l'ensemble contenant les entiers 1, 2, 3 et 4.



Il est possible de définir un ensemble de deux manières :

- **paramétrique**, comme par exemple $\{n^2, n \in \mathbb{N}\}$;
- **implicite**, comme par exemple $\{x \in \mathbb{R}, x^3 - 2x + 1 = 0\}$.

Définition 3.2

Soient E et F deux ensembles. Nous disons que E est **inclus** dans F et nous notons $E \subset F$, si et seulement si tous les éléments de E appartiennent aussi à F . Nous disons aussi que E est une partie de F ou que F contient E . L'ensemble des parties de F se note $\mathcal{P}(F)$.

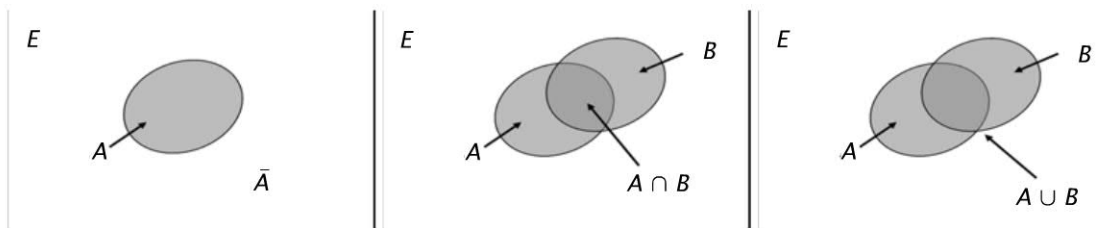
Exemples

1. L'ensemble des entiers naturels $\mathbb{N}$ est contenu dans l'ensemble des réels $\mathbb{R}$.
2. $\{1\}$ est contenu dans $\{1; 2; 3; 4\}$.

Définition 3.3

Soient E un ensemble, A et B des parties de E .

- Nous définissons le **complémentaire** de A dans E par $\bar{A} = \{x \in E; x \notin A\}$.
- Nous définissons l'**intersection** de A et B par $A \cap B = \{x \in E; x \in A \text{ et } x \in B\}$.
- Nous définissons l'**union** de A et B par $A \cup B = \{x \in E; x \in A \text{ ou } x \in B\}$.
- Deux ensembles A et B sont **disjoints** s'il n'existe aucun élément commun à A et à B , c'est-à-dire si $A \cap B = \emptyset$.

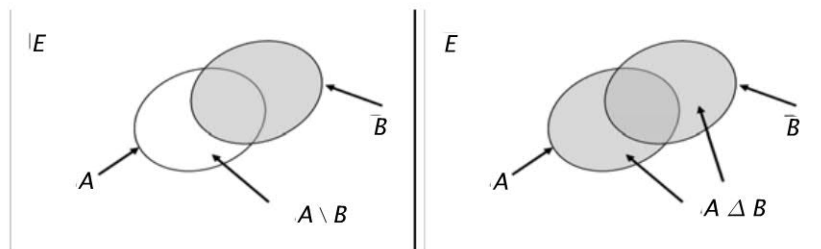


Le « ou » dans la définition de l'union a un sens inclusif, c'est-à-dire que $A \cup B$ est l'ensemble des éléments x de E qui appartiennent à l'une au moins des parties A et B .

Définition 3.4

Soient E un ensemble, A et B des parties de E . Nous définissons

- la **différence** par $A \setminus B = \{x \in E; x \in A \text{ et } x \notin B\} = A \cap \bar{B}$.
- la **différence symétrique** par $A \Delta B = (A \cup B) \setminus (A \cap B) = (A \cap \bar{B}) \cup (\bar{A} \cap B)$.



$A \Delta B$ est l'ensemble des éléments qui appartiennent à une et une seule des parties A et B .

Définition 3.5

- Un **recouvrement** d'une partie A de E est une famille $(A_i)_{i \in I}$ de parties telles que $A_i \subset E$ et $A \subset \bigcup_{i \in I} A_i$.
- Une **partition** P d'un ensemble E est une famille de parties non vides de E $(A_i)_{i \in I}$ telles que $A_i \cap A_j = \emptyset$ si $i \neq j$ et $\bigcup_{i \in I} A_i = E$.

Définition 3.6

Soient A et B deux ensembles. Alors le **produit cartésien** de ces deux ensembles, noté $A \times B$, est l'ensemble des couples $(a; b)$ tels que $a \in A$ et $b \in B$.

$$A \times B = \{(a; b) ; x \in A \text{ et } y \in B\}.$$

Exemple

$$\mathbb{R}^2 = \mathbb{R} \times \mathbb{R} = \{(x; y); x \in \mathbb{R} \text{ et } y \in \mathbb{R}\}.$$



Plus généralement, le produit cartésien de n ensembles E_i est :

$$E_1 \times \dots \times E_n = \{(x_1, \dots, x_n) ; x_1 \in E_1 ; \dots ; x_n \in E_n\}.$$

Si $E_1 = E_2 = \dots = E_n = E$, alors nous notons E^n .

2. Propriétés

Théorème 3.1

Soient A, B et C des parties d'un ensemble E . Alors, nous avons les propriétés suivantes sur le complémentaire, la réunion et l'intersection :

- $\overline{\overline{E}} = E$ et $\overline{\emptyset} = E$,
- $\overline{\overline{A}} = A$,
- si $A \subset B$, alors $\overline{B} \subset \overline{A}$,
- $\overline{A \cap B} = \overline{A} \cup \overline{B}$ et $\overline{A \cup B} = \overline{A} \cap \overline{B}$ (lois de Morgan),
- $A \cup B = B \cup A$: commutativité de l'union,
- $A \cap B = B \cap A$: commutativité de l'intersection,
- $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$: associativité de l'union,
- $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$: associativité de l'intersection,
- $A \cup A = A$; $A \cup \emptyset = A$ et $A \cup E = E$,
- $A \cap A = A$; $A \cap \emptyset = \emptyset$ et $A \cap E = A$,
- $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$: distributivité de l'intersection par rapport à l'union,
- $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$: distributivité de l'union par rapport à l'intersection.

Application

Nous allons montrer trois des propriétés précédemment énoncées ci-dessus. Cela va nous permettre d'utiliser le raisonnement par double inclusion, très utile pour effectuer des démonstrations.

1. Montrer que $\overline{\overline{A}} = A$.

2. Montrer que $\overline{A \cap B} = \overline{A} \cup \overline{B}$.

3. Montrer que $\overline{A \cup B} = \overline{A} \cap \overline{B}$.

1.



Pour cela, nous devons montrer la double inclusion $\overline{\overline{A}} \subset A$ et $A \subset \overline{\overline{A}}$.

Soit $x \in \overline{\overline{A}}$, alors $x \notin \overline{A}$ et donc $x \in A$. Nous avons bien $\overline{\overline{A}} \subset A$.

Réciproquement : soit $x \in A$, alors $x \notin \overline{A}$ et donc $x \in \overline{\overline{A}}$. Nous avons bien $A \subset \overline{\overline{A}}$.

La double inclusion permet donc de conclure que $\overline{\overline{A}} = A$.

2.



Nous devons montrer la double inclusion $\overline{A \cap B} \subset \overline{A} \cup \overline{B}$ et $\overline{A} \cup \overline{B} \subset \overline{A \cap B}$.

Soit $x \in \overline{A \cap B}$, alors $x \notin A \cap B$. Nous avons donc en particulier $x \notin A$ ou $x \notin B$ c'est-à-dire $x \in \overline{A}$ ou $x \in \overline{B}$. Donc $x \in \overline{A} \cup \overline{B}$ et $\overline{A \cap B} \subset \overline{A} \cup \overline{B}$.

Réciproquement : soit $x \in \overline{A} \cup \overline{B}$, alors, $x \in \overline{A}$ ou $x \in \overline{B}$ c'est-à-dire que $x \notin A$ ou $x \notin B$. En particulier, $x \notin A \cap B$ donc $x \in \overline{A \cap B}$. Nous avons donc $\overline{A} \cup \overline{B} \subset \overline{A \cap B}$.

La double inclusion permet de conclure que $\overline{A \cap B} = \overline{A} \cup \overline{B}$.

3.



Nous devons pour cela montrer la double inclusion $\overline{A \cup B} \subset \overline{A} \cap \overline{B}$ et $\overline{A} \cap \overline{B} \subset \overline{A \cup B}$.

Soit $x \in \overline{A \cup B}$, alors $x \notin A \cup B$. Nous avons donc en particulier $x \notin A$ et $x \notin B$ c'est-à-dire $x \in \overline{A}$ et $x \in \overline{B}$. Donc $x \in \overline{A} \cap \overline{B}$ et ainsi nous avons montré que $\overline{A \cup B} \subset \overline{A} \cap \overline{B}$.

Réciproquement : soit $x \in \overline{A} \cap \overline{B}$, alors $x \in \overline{A}$ et $x \in \overline{B}$ c'est-à-dire que $x \notin A$ et $x \notin B$. Donc $x \notin A \cup B$ et $x \in \overline{A \cup B}$. Nous avons montré $\overline{A} \cap \overline{B} \subset \overline{A \cup B}$.

Cela nous permet de conclure que $\overline{A \cup B} = \overline{A} \cap \overline{B}$.



Nous aurions également pu raisonner par équivalence pour réaliser ces démonstrations mais lorsque les preuves à effectuer sont plus complexes, cela est souvent source d'erreurs et il vaut mieux privilégier le raisonnement par double inclusion.

Pour s'entraîner (solutions p. 506)

3.1 Soient $A = \{1; 2; 3\}$ et $B = \{2; 3; 4\}$. Déterminer $A \cap B$ et $A \cup B$.

3.2 Soit $A = \{x \in \mathbb{R}; x > 9\}$. Déterminer le complémentaire $\overline{A}$ de A .

3.3 Trouver un exemple d'ensembles A , B et C tels que $A \cup B = A \cup C$ et $B \neq C$.

3.4 Démontrer que si $A \cup B = A \cap B$ alors $A = B$.

3.5 Démontrer que si $A \cup B = A \cup C$ et $A \cap B = A \cap C$ alors $B = C$.

3.6 Simplifier les quatre expressions suivantes :

1. $\overline{\overline{A \cup B}}$,

2. $\overline{\overline{A \cap B}}$,

3. $\overline{A} \cup (A \cap B)$,

4. $\overline{A} \cap (A \cup \overline{B})$.

Définition 4.1

Soient E et F deux ensembles.

- Une **relation binaire**, ou correspondance, de E vers F est la donnée du triplet (E, Γ, F) où Γ est une partie de $E \times F$. Pour $x \in E$ et $y \in F$, lorsque $(x, y) \in \Gamma$, nous notons cette relation par $x \mathcal{R} y$.
 E est l'**ensemble de départ** de $\mathcal{R}$, F est l'**ensemble d'arrivée** de $\mathcal{R}$ et Γ est le **graphe** de la relation $\mathcal{R}$.
- Une **relation binaire** $\mathcal{R}$ sur E est définie par le choix d'une partie Γ de $E \times E$.

Exemples

1. La divisibilité sur $\mathbb{N}^*$, définie en 9.1, est une relation binaire.
2. L'inclusion sur $\mathcal{P}(E)$, les parties de l'ensemble E , est une relation binaire.
3. Les relations de congruence sur $\mathbb{N}^*$, définies en 9.6, sont également des relations binaires.

Définition 4.2

Si F est une partie de E , la **relation induite** par $\mathcal{R}$ sur E est simplement la relation $x \mathcal{R} y$ entre éléments de F .

Définition 4.3

Une relation binaire $\mathcal{R}$ définie sur un ensemble E est :

- **réflexive** si elle vérifie $\forall x \in E, x \mathcal{R} x$,
- **symétrique** si $\forall x \in E, \forall y \in E, x \mathcal{R} y \implies y \mathcal{R} x$,
- **antisymétrique** si elle vérifie $\forall x \in E, \forall y \in E, (x \mathcal{R} y \text{ et } y \mathcal{R} x) \implies x = y$,
- **transitive** si elle vérifie $\forall x \in E, \forall y \in E, \forall z \in E, (x \mathcal{R} y \text{ et } y \mathcal{R} z) \implies x \mathcal{R} z$.

Exemple

Dans $\mathbb{N}$, la relation $\leq$ est réflexive, antisymétrique et transitive. Elle vérifie en effet,

- $\forall x \in \mathbb{N}, x \leq x$ donc elle est réflexive.
- $\forall x \in \mathbb{N}, \forall y \in \mathbb{N}, x \leq y \text{ et } y \leq x \iff x = y$, donc elle est antisymétrique.
- $\forall x \in \mathbb{N}, \forall y \in \mathbb{N}, \forall z \in \mathbb{N}, x \leq y, y \leq z \iff x \leq z$, donc elle est transitive.

Définition 4.4

- Une relation binaire $\mathcal{R}$ définie sur un ensemble E est une **relation d'équivalence** si elle est à la fois réflexive, symétrique et transitive.
- Si $x \in E$, nous appelons **classe d'équivalence** de x modulo $\mathcal{R}$ l'ensemble des y de E tels que $x \mathcal{R} y$.

Exemples

1. La relation d'égalité sur n'importe quel ensemble E est une relation d'équivalence et c'est la plus forte (elle est contenue dans toutes les autres.)
2. La relation d'équivalence la plus faible est celle dont le graphe est $E \times E$, c'est-à-dire celle qui est satisfaite par tous les couples $(x; y) \in E \times E$.

Définition 4.5

Une relation binaire $\mathcal{R}$ définie dans un ensemble E est une **relation d'ordre** si elle est à la fois réflexive, antisymétrique et transitive. Nous la notons souvent $<$.

Exemples

1. La relation $\leq$ est une relation d'ordre.
2. La divisibilité sur $\mathbb{N}$ est une relation d'ordre.



La seule relation à la fois d'ordre et d'équivalence est l'égalité.

Définition 4.6

- Deux éléments x et y d'un ensemble E sont dits **comparables** pour $<$ si et seulement si $x < y$ ou $y < x$,
- Une relation d'ordre $<$ dans E est dite d'**ordre total** si deux éléments quelconques x et y de E sont toujours comparables, c'est-à-dire si quels que soient x et $y \in E$, nous avons $x < y$ ou $y < x$. Dans le cas contraire l'ordre est **partiel**.

Exemple

Dans $\mathbb{R}$, la relation $\leq$ est d'ordre total.

Définition 4.7

Soit A une partie d'un ensemble E .

- S'il existe un élément a de E tel que, pour tout $x \in A$ nous ayons $x < a$, nous disons que a est un **majorant** de A et que A est une partie majorée de E .
- S'il existe un élément b de E tel que, pour tout $x \in A$ nous ayons $x > b$, nous disons que b est un **minorant** de A et que A est une partie minorée de E .
- Une partie **bornée** de E est une partie qui est à la fois majorée et minorée.

Définition 4.8

Soit A une partie d'un ensemble E .

- Un élément a de E est appelé **plus grand élément** de A si $a \in A$ et si a est un majorant de A . Si un tel élément existe il est unique.
- Un élément b de E est appelé **plus petit élément** de A si $b \in A$ et si b est un minorant de A . Si un tel élément existe il est unique.



Il est important de bien faire la distinction entre majorant et plus grand élément d'une partie A d'un ensemble E . Contrairement au majorant de A qui n'appartient pas nécessairement à A , le plus grand élément de A est un élément appartenant à A . Cette remarque s'applique également pour distinguer minorant et plus petit élément d'une partie A d'un ensemble E .

Définition 4.9

Soit A une partie d'un ensemble E .

- Nous appelons **borne supérieure** d'une partie majorée A , le plus petit des majorants de A . Si cette borne existe elle est unique. Elle est notée $\sup_E(A)$.
- Nous appelons **borne inférieure** d'une partie minorée A , le plus grand des mineurs de A . Si cette borne existe elle est unique. Elle est notée $\inf_E(A)$.

Application

Soit $\mathcal{R}$ la relation définie sur $\mathbb{R}$ par

$$x\mathcal{R}y \iff x^3 - y^3 = 3(x - y).$$

Montrer qu'il s'agit d'une relation d'équivalence et déterminer les classes d'équivalence de cette relation.



Montrons d'abord qu'il s'agit d'une relation d'équivalence.

- Tout d'abord la relation est réflexive. En effet, $\forall x \in \mathbb{R}$,

$$x^3 - x^3 = 3(x - x) = 0.$$

Donc $x\mathcal{R}x$.

- Elle est également symétrique. En effet, $\forall x, y \in \mathbb{R}$,

$$\begin{aligned} x^3 - y^3 = 3(x - y) &\implies -(x^3 - y^3) = -3(x - y), \\ &\implies y^3 - x^3 = 3(y - x). \end{aligned}$$

Donc $x\mathcal{R}y \implies y\mathcal{R}x$.

- Enfin, elle est transitive. En effet, $\forall x, y, z \in \mathbb{R}$,

$$\begin{aligned} (x\mathcal{R}y \text{ et } y\mathcal{R}z) &\implies x^3 - z^3 = x^3 - y^3 + y^3 - z^3 \\ &= 3(x - y) + 3(y - z) = 3(x - z). \end{aligned}$$

Donc $x\mathcal{R}z$.

Nous en déduisons donc que la relation $\mathcal{R}$ est une relation d'équivalence.



Cherchons à présent les classes d'équivalence pour cette relation, c'est-à-dire l'ensemble des couples $(x, y) \in \mathbb{R}^2$ tels que $x\mathcal{R}y$. Pour cela, résolvons l'équation $x^3 - y^3 = 3(x - y)$.

$$\begin{aligned} x^3 - y^3 &= 3(x - y) \\ \implies x^3 - y^3 - 3(x - y) &= 0 \\ \implies (x - y)(x^2 + xy + y^2) - 3(x - y) &= 0 \\ \implies (x - y)(x^2 + xy + y^2 - 3) &= 0. \end{aligned}$$

Les solutions sont telles que $x - y = 0$ ou $x^2 + xy + y^2 - 3 = 0$.

Les classes d'équivalence pour la relation $\mathcal{R}$ sont donc les $y \in \mathbb{R}$ tels que $x = y$ ou $x^2 + xy + y^2 - 3 = 0$.

Pour s'entraîner (solutions p. 506)

4.1 Nous définissons la relation $<$ par $x < y$ si et seulement si $x \leq y$ et $x \neq y$. Montrer que ce n'est pas une relation d'ordre.

4.2 Nous considérons les relations suivantes :

1. $E = \mathbb{N}, x\mathcal{R}y \iff x = -y,$

2. $E = \mathbb{R}, x\mathcal{R}y \iff \cos^2 x + \sin^2 y = 1,$

3. $E = \mathbb{N}, x\mathcal{R}y \iff \exists p, q \geq 1, y = px^q$ avec p et q entiers.

Ces relations sont-elles réflexives, symétriques, transitives ? Sont-elles des relations d'ordre, d'équivalence ?

4.3 Nous considérons la relation suivante : $x\mathcal{R}y$ si et seulement si $x^2 - y^2 = x - y$. Montrer que c'est une relation d'équivalence.

Déterminer la classe d'équivalence d'un élément $x \in \mathbb{R}$. Combien y-a-t-il d'éléments dans cette classe ?

4.4 Soient a et b deux réels strictement positifs. Les parties suivantes sont-elles majorées, minorées ? Si oui, donner les bornes supérieures et/ou inférieures.

1. $\{a + bn; n \in \mathbb{N}\},$

2. $\{a + \frac{b}{n}; n \in \mathbb{N}^*\},$

3. $\{a + (-1)^n b; n \in \mathbb{N}\}.$

4.5 Soient A et B deux parties non vides de $\mathbb{R}$ avec $\forall a \in A, \forall b \in B, a \leq b$. Montrer que A est majorée, B est minorée et que $\sup(A) \leq \inf(B)$.

Définition 5.1

Une **fonction** f de l'ensemble de départ E dans l'ensemble d'arrivée F est définie par une correspondance qui permet d'associer à tout $x \in E$ **au plus un élément** $y \in F$. Nous appelons **ensemble de définition** de f l'ensemble des éléments x de E tels que $\exists y \in F$ tel que $y = f(x)$. Pour tout $x \in E$, l'élément $y \in F$ tel que $y = f(x)$, s'il existe, est appelé l'**image** de x par f . Pour tout élément $y \in F$, tout élément $x \in E$ tel que $y = f(x)$ est un **antécédent** de y par f . Les fonctions de E dans F forment un ensemble noté $\mathcal{F}(E, F)$.



Un élément x a au plus une image. En revanche, un élément y peut avoir un nombre quelconque d'antécédents. Par exemple, soit la fonction $x \mapsto \cos x$; 1 a une infinité d'antécédents et 2 n'en a aucun.

Définition 5.2

Une application f de l'ensemble de départ E dans l'ensemble d'arrivée F est définie par une correspondance qui permet d'associer à tout $x \in E$ **un unique élément** $y \in F$ que nous notons $f(x)$. Pour tout $x \in E$, l'élément $y \in F$ tel que $y = f(x)$ est appelé l'**image** de x par f . Pour tout élément $y \in F$, tout élément $x \in E$ tel que $y = f(x)$ est un **antécédent** de y par f . Les applications de E dans F forment un ensemble noté $\mathcal{F}(E, F)$.



Ainsi, une application fournit une image à tout élément de E , quand une fonction ne le fait qu'au sein de son domaine de définition.

Exemples

1. L'application de E dans E qui à x associe x est l'application identité. Nous la notons Id_E .
2. L'application de $\mathbb{R}$ dans $\mathbb{R}$ qui à x associe $\sqrt{x}$ est une fonction dont le domaine de définition est $\mathbb{R}_+$.

Définition 5.3

Soient E, F et G des ensembles. Soit f une application de E dans G et g une application de F dans G . Si $E \subset F$ et si, pour tout $x \in E$, nous avons $f(x) = g(x)$, nous disons que f est une **restriction** de g ou que g est un **prolongement** de f .

Définition 5.4

Soient E, F et G trois ensembles, f une application de E dans F et g une application de F dans G . La **composée** de f par g est l'application de E dans G notée par $g \circ f$ et définie par :

$$x \mapsto (g \circ f)(x) = g(f(x)).$$

Théorème 5.1

La loi de composition des applications est **associative**. En effet, si nous considérons E, F, G et H des ensembles et $f : E \longrightarrow F, g : F \longrightarrow G$ et $h : G \longrightarrow H$ des applications,

$$(h \circ g) \circ f = h \circ (g \circ f) = h \circ g \circ f.$$

Ainsi lorsque nous composons des applications, il n'est pas nécessaire de préciser l'ordre dans lequel il faut les appliquer et nous pouvons donc supprimer les parenthèses.

Définition 5.5

Une application f de E dans F est dite **injective** (ou est une **injection**) si elle vérifie l'une des deux propriétés équivalentes suivantes :

$$\forall x \in E, \forall x' \in E, x \neq x' \implies f(x) \neq f(x'),$$

$$\forall x \in E, \forall x' \in E, f(x) = f(x') \implies x = x'.$$

Définition 5.6

Une application f de E dans F est dite **surjective** (ou est une **surjection**) si tout élément y de F est l'image d'au moins un élément x de E soit :

$$\forall y \in F, \exists x \in E, y = f(x).$$

Définition 5.7

Une application f de E dans F est dite **bijective** (ou est une **bijection**) si elle est à la fois injective et surjective. Dans ce cas, tout élément y de F est l'image d'un unique élément x de E , noté $f^{-1}(y)$. L'application f^{-1} est la bijection réciproque de f . Nous avons donc :

$$x = f^{-1}(y) \iff y = f(x).$$



La définition précédente entraîne que $f \circ f^{-1} = Id_F$ et $f^{-1} \circ f = Id_E$.

Théorème 5.2

Soient E, F et G des ensembles, f une application de E dans F et g une application de F dans G . Alors,

- Si f et g sont injectives, $g \circ f$ est injective,
- Si $g \circ f$ est injective, f est injective,
- Si f et g sont surjectives, $g \circ f$ est surjective,
- Si $g \circ f$ est surjective, g est surjective,
- Si f et g sont bijectives, $g \circ f$ est bijective et $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

Définition 5.8

Soit f une application de E dans F . Si $A \subset E$, nous appelons **image** de A par f la partie de F constituée par les images des éléments de A :

$$f(A) = \{y \in F; \exists x \in A, f(x) = y\}.$$

Définition 5.9

Soit f une application de E dans F . Si $B \subset F$, nous appelons **image réciproque** de B la partie de E constituée par les éléments dont l'image est dans B :

$$f^{-1}(B) = \{x \in E; f(x) \in B\}.$$

Théorème 5.3

Soient A_1, A_2, B_1, B_2 des ensembles. Alors nous avons :

$$A_1 \subset A_2 \implies f(A_1) \subset f(A_2),$$

$$B_1 \subset B_2 \implies f^{-1}(B_1) \subset f^{-1}(B_2),$$

$$f(A_1 \cup A_2) = f(A_1) \cup f(A_2),$$

$$f(A_1 \cap A_2) \subset f(A_1) \cap f(A_2),$$

$$f^{-1}(B_1 \cup B_2) = f^{-1}(B_1) \cup f^{-1}(B_2),$$

$$f^{-1}(B_1 \cap B_2) = f^{-1}(B_1) \cap f^{-1}(B_2).$$

Application

1. Montrer que l'application $f(x) = x^3$ est bijective de $\mathbb{R}$ dans $\mathbb{R}$.
2. Montrer que l'application $f(x) = x^2$ n'est pas bijective de $\mathbb{R}$ dans $\mathbb{R}$.

1.



Tout d'abord, f est injective.

En effet, soient x et $y \in \mathbb{R}$ tels que $f(x) = f(y)$. Nous allons montrer que $x = y$.

Nous avons $f(x) = f(y) \iff x^3 = y^3 \iff x^3 - y^3 = 0 \iff (x - y)(x^2 + xy + y^2) = 0$. Comme $x^2 + xy + y^2$ n'admet pas de racines réelles, la seule solution est que $x - y = 0$ ou encore $x = y$ ce qui prouve bien l'injectivité de l'application.



D'autre part, f est surjective.

En effet, nous avons $f(\mathbb{R}) = \mathbb{R}$. Cela permet de dire que $\forall y \in \mathbb{R}, \exists x \in \mathbb{R}$ tel que $y = f(x)$. L'application f est donc surjective.

Le fait que f soit à la fois injective et surjective nous permet de conclure que f est bijective.

2.

Tout d'abord, f n'est pas injective. En effet, soient x et $y \in \mathbb{R}$ tels que $f(x) = f(y)$. Nous allons montrer que nous n'avons pas uniquement $x = y$ comme solution. Nous avons $f(x) = f(y) \iff x^2 = y^2 \iff x^2 - y^2 = 0 \iff (x - y)(x + y) = 0$. Les solutions sont $x - y = 0$ ou $x + y = 0$ soit $x = y$ ou $x = -y$. Nous en déduisons donc la non-injectivité de la relation, ce qui suffit à montrer que l'application n'est pas bijective.



Nous aurions pu également montrer que f n'est pas surjective. En effet, nous avons que $f(\mathbb{R}) = \mathbb{R}_+$, ce qui suffit à montrer la non surjectivité dans $\mathbb{R}$. L'application n'est pas bijective.

En revanche, si nous étudions la même application $f(x) = x^2$ mais cette fois de $\mathbb{R}_+$ dans $\mathbb{R}_+$, nous pouvons montrer que dans ce cas l'application est bijective. Elle est déjà surjective puisque $f(\mathbb{R}_+) = \mathbb{R}_+$. En reprenant le raisonnement précédent concernant l'éventuelle injectivité de l'application, à l'étape $x - y = 0$ ou $x + y = 0$ nous ne conservons que $x - y = 0$ ou $x + y = 0 \implies x = y = 0$ puisque nous avons supposé que x et y sont positifs. Dans les deux cas, nous déduisons $x = y$. Nous en déduisons que l'application est injective et donc bijective.

Pour s'entraîner (solutions p. 507)

5.1 Nous considérons les trois applications suivantes $f_1 : x \mapsto x^2$, $f_2 : x \mapsto \sqrt{x}$ et $f_3 : x \mapsto x - 4$. Calculer, lorsque cela est possible, $f_3 \circ f_1$, $f_1 \circ f_3$, $f_3 \circ f_2$, $f_2 \circ f_3$ et $f_1 \circ f_1$.

5.2 Nous considérons les deux applications suivantes définies de $\mathbb{R}$ dans $\mathbb{R}$ par $f : x \mapsto 3x + 1$ et $g : x \mapsto x^2 - 1$. Calculer $f \circ g$ et $g \circ f$. Avons-nous l'égalité entre ces deux quantités ?

5.3 Nous considérons l'application définie de $[1; +\infty[$ dans $[0; +\infty[$ par $f : x \mapsto x^2 - 1$. L'application f est-elle bijective ?

5.4 Nous considérons l'application définie de $\mathbb{N}$ dans $\mathbb{N}$ par $f : n \mapsto n + 1$. L'application f est-elle injective, surjective, bijective ?

5.5 Nous considérons l'application définie de $\mathbb{R}^2$ dans $\mathbb{R}^2$ par $f : (x; y) \mapsto (x + y, x - y)$. L'application f est-elle injective, surjective, bijective ?

5.6 Soient A, B et C des ensembles et $f : A \longrightarrow B$ et $g : B \longrightarrow C$ des applications.

1. Montrer que $g \circ f$ injective $\implies f$ injective.
2. Montrer que $g \circ f$ surjective $\implies g$ surjective.

5.7 Nous considérons l'application tangente définie sur $\left]-\frac{\pi}{2}; \frac{\pi}{2}\right[$. Montrer qu'elle admet une bijection réciproque $\tan^{-1}$ et donner ses caractéristiques.

1. Entiers naturels

L'ensemble des **entiers naturels** est l'ensemble des nombres le plus simple à savoir $\{0; 1; 2; \dots\}$. Il est d'usage de le noter $\mathbb{N}$ et sa construction ne sera pas donnée avec plus de détails dans ce livre.

Théorème 6.1

L'ensemble $\mathbb{N}$ des entiers naturels est totalement ordonné et vérifie les trois propriétés suivantes :

- toute partie non vide de $\mathbb{N}$ a un plus petit élément,
- toute partie non vide et majorée de $\mathbb{N}$ a un plus grand élément,
- $\mathbb{N}$ n'a pas de plus grand élément.



D'autres ensembles faisant intervenir des nombres entiers peuvent être définis.

$\mathbb{Z} = \{\dots; -2; -1; 0; 1; 2; \dots\}$ est l'ensemble des **entiers relatifs**.

Soit $\mathbb{A}$ l'ensemble des classes d'équivalence de $\mathbb{N} \times \mathbb{N}$ par la relation d'équivalence $\mathcal{R}$ définie par $(n_1; n'_1) \mathcal{R} (n_2; n'_2) \Leftrightarrow (n_1 + n'_2 = n'_1 + n_2)$. Nous remarquons que l'addition usuelle sur $\mathbb{N} \times \mathbb{N}$, ainsi que le produit défini par $(n_1; n'_1) * (n_2; n'_2) = (n_1 n_2 + n'_1 + n'_2; n_1 n'_2 + n'_1 n_2)$, sont compatibles avec la relation $\mathcal{R}$ et peuvent donc être définies sur l'ensemble $\mathbb{A}$. Puis, chaque couple $(n; n')$, se trouve dans l'une des classes $(d; 0)$ si $n > n'$, $(0; d)$ si $n < n'$ ou $(0; 0)$ si $n = n'$, ce qui permet de retrouver la forme classique de $\mathbb{Z}$. Enfin, les opérations usuelles d'addition et de multiplication définies sur $\mathbb{A}$ coïncident avec celles sur $\mathbb{Z}$.

$\mathbb{Q}$ est l'ensemble des **nombres rationnels** c'est-à-dire des fractions de nombres entiers.

Il peut également être défini à l'aide de la relation $\mathcal{R}$ définie sur $\mathbb{Z} \times \mathbb{Z}^*$ par $(a; b) \mathcal{R} (a'; b') \Leftrightarrow ab' = ba'$. $\mathcal{R}$ est une relation d'équivalence et un nombre rationnel est la classe d'équivalence d'un élément $(a; b) \in \mathbb{Z} \times \mathbb{Z}^*$. Il est noté $\frac{a}{b}$. L'ensemble des classes d'équivalences $\mathbb{Z} \times \mathbb{Z}^*$ par $\mathcal{R}$ correspond à $\mathbb{Q}$. L'addition et le produit coordonnée par coordonnée sont des opérations compatibles avec $\mathcal{R}$ et coïncident avec les opérations usuelles sur $\mathbb{Q}$.

Le **raisonnement par récurrence** s'appuie directement sur les propriétés de l'ensemble $\mathbb{N}$ des nombres entiers naturels, c'est pourquoi nous le mentionnons à nouveau.

Définition 6.1 Raisonement par récurrence

Soit $n_0 \in \mathbb{N}$. Soit $E(n)$ un énoncé qui porte sur un entier naturel $n \geq n_0$. Pour que $E(n)$ soit vrai $\forall n \in \mathbb{N}$ tel que $n \geq n_0$, il faut que $E(n_0)$ soit vrai et que $\forall n \in \mathbb{N}, n \geq n_0$, l'implication $E(n) \Rightarrow E(n+1)$ soit vraie.

Définition 6.2

Soit $(a; b) \in \mathbb{N}^2$. a **divise** b dans $\mathbb{N}$, noté $a|b$, si et seulement si $\exists c \in \mathbb{N}, b = ac$.

Définition 6.3

Un nombre $p \in \mathbb{N}$ est **premier** si et seulement si $p \geq 2$ et $\forall a \in \mathbb{N}, a|p \implies a = 1$ ou $a = p$. Autrement dit, p est premier si et seulement s'il est supérieur ou égal à 2 et ses seuls diviseurs sont 1 et lui-même.

Théorème 6.2

- Il y a une infinité de nombres premiers.
- Si n n'est divisible par aucun nombre premier inférieur ou égal à $\sqrt{n}$, alors il est premier.
- Tout entier n , avec $n \geq 2$, s'écrit de façon unique comme produit, à l'ordre des facteurs près, de nombres premiers :

$$n = \prod_{k=1}^K p_i^{\alpha_i}$$

où les nombres entiers p_i sont distincts et premiers deux à deux et les α_i appartiennent à $\mathbb{N}^*$.

2. Ensembles finis, infinis**Définition 6.4**

Un ensemble E est **équipotent** à un ensemble F si et seulement s'il existe une bijection de E sur F .

Exemple

$\mathbb{N}$ et $\mathbb{Q}$ sont équipotents.

Définition 6.5

- Un ensemble E non vide est **fini** s'il existe $n \in \mathbb{N}$ tel que E soit équipotent à $F_n = \{1; \dots; n\}$. Le nombre n est le **cardinal** (ou nombre d'éléments) de E . Nous le notons $n = \text{Card}(E)$. Nous convenons que l'ensemble vide est fini et que $\text{Card}(\emptyset) = 0$.

Proposition 6.1

Soit E un ensemble fini. Toute partie A de E est finie et nous avons $\text{Card}(A) \leq \text{Card}(E)$. L'égalité des cardinaux a lieu si et seulement si $A = E$.

Théorème 6.3

Soient E et F deux ensembles finis de même cardinal et f une application de E dans F . Nous avons l'équivalence des trois propriétés :

$$f \text{ bijective} \iff f \text{ injective} \iff f \text{ surjective.}$$



Dans ce cas, pour démontrer que f est bijective, il suffit de démontrer que f est injective ou f surjective.

Théorème 6.4

Soient E et F deux ensembles finis. Alors $E \cup F$ est fini et

$$\text{Card}(E \cup F) + \text{Card}(E \cap F) = \text{Card}(E) + \text{Card}(F).$$

Corollaire 6.1

- Soient E un ensemble fini et F une partie de E . Si $\text{Card}(F) = \text{Card}(E)$ alors $F = E$.
- Soient $n \in \mathbb{N}^*$ et $E_1, \dots, E_n$ des ensembles finis. Si $E_1, \dots, E_n$ sont deux à deux dis-joints, alors $\text{Card}(E_i) = \sum_{i=1}^n \text{Card}(E_i)$.

Définition 6.6

Un ensemble est **infini** s'il n'est pas fini.

Exemple

$\mathbb{N}$ est infini.

Théorème 6.5

Tout ensemble admettant au moins une partie infinie est lui-même infini.

Application

1. Montrer que $\forall n \in \mathbb{N}^*, 4^n + 5$ est un multiple de 3.

2. Montrer que $\forall n \in \mathbb{N}^*, \sum_{k=1}^n \frac{1}{k(k+1)} = 1 - \frac{1}{n+1}$.

1.



Nous allons utiliser pour cela le raisonnement par récurrence qui est, comme nous l'avons déjà indiqué dans une précédente Fiche, très utilisé dans de nombreuses démonstrations mathématiques.

- Si $n = 1$, $4^n + 5 = 4 + 5 = 9$ qui est bien un multiple de 3.
- Supposons que cette propriété soit vraie à l'ordre n fixé. Montrons qu'elle reste vraie à l'ordre $n + 1$.

$$4^{n+1} + 5 = 4^n \times 4 + 5 = 4^n + 5 + 3 \times 4^n.$$

De plus, par hypothèse de récurrence, $4^n + 5$ est un multiple de 3 et il est évident que 3×4^n est également un multiple de 3. Nous en déduisons que $4^{n+1} + 5$ est un multiple de 3 et donc la propriété est vraie à l'ordre $n + 1$.

Nous pouvons donc conclure que la propriété est vraie pour tout $n \in \mathbb{N}^*$.

2.



Là encore, nous allons utiliser le raisonnement par récurrence pour démontrer cette égalité.

- Pour $n = 1$, $\frac{1}{2} = 1 - \frac{1}{1+1} = 1 - \frac{1}{2} = \frac{1}{2}$, donc la propriété est vraie à l'ordre 1.
- Supposons la propriété vraie à l'ordre n fixé et montrons alors qu'elle reste vraie à l'ordre $n + 1$. En utilisant l'hypothèse de récurrence, nous obtenons :

$$\begin{aligned}
 \sum_{k=1}^{n+1} \frac{1}{k(k+1)} &= \sum_{k=1}^n \frac{1}{k(k+1)} + \frac{1}{(n+1)(n+2)} \\
 &= 1 - \frac{1}{n+1} + \frac{1}{(n+1)(n+2)} \\
 &= 1 - \frac{(n+1)}{(n+1)(n+2)} \\
 &= 1 - \frac{1}{(n+2)}.
 \end{aligned}$$

La propriété est donc vraie à l'ordre $n + 1$.

Nous pouvons donc conclure que la propriété est vraie pour tout $n \in \mathbb{N}^*$.

Pour s'entraîner (solutions p. 507)

- | | |
|---|---|
| <p>6.1 Montrer que $\forall n \geq 4$ entier, $2^n < n!$.</p> <p>6.2 Montrer que $\forall n \in \mathbb{N}^*$, $3 \times 5^{2n-1} + 2^{3n-2}$ est divisible par 17.</p> <p>6.3 Montrer par récurrence que $\forall n \in \mathbb{N}^*$,
 $1 + 2 + \dots + n = \frac{n(n+1)}{2}$.</p> <p>6.4 Montrer par récurrence que $\forall n \in \mathbb{N}^*$,
 $1^2 + 2^2 + \dots + n^2 = \frac{n(n+1)(2n+1)}{6}$.</p> <p>6.5 Montrer qu'aucun des entiers $n! + 2, \dots, n! + n$ n'est premier.</p> | <p>6.6 Décomposer 46848 en facteurs premiers.</p> <p>6.7 Quel est le cardinal de l'ensemble $E = \{2; 6; 7; 10\}$?</p> <p>6.8 Soient A et B deux parties des ensembles E et F respectivement et f une application de E dans F. Montrer que si $f(A)$ est une partie finie de F, alors A n'est pas nécessairement une partie finie de E.</p> |
|---|---|

Structures algébriques : groupes

1. Loi de composition interne

Définition 7.1

Une **loi de composition interne**, notée $*$, sur un ensemble E est une application de $E \times E$ dans E . À un couple $(x; y)$ de $E \times E$, nous associons un élément de E , noté $x * y$. L'élément $x * y$ s'appelle composé de x et de y .

Exemple

Dans $\mathbb{R}$, nous pouvons citer la loi de composition interne '+'.

Définition 7.2

Une loi de composition interne $*$ sur E :

- est **associative** si $\forall x, y, z \in E, (x * y) * z = x * (y * z)$,
- est **commutative** si $\forall x, y \in E, x * y = y * x$,
- admet un **élément neutre** si $\exists e \in E$ tel que $\forall x \in E, x * e = e * x = x$.

Exemples

1. Les lois de composition interne '+' et '×' sur $\mathbb{C}$ sont associatives et commutatives et admettent pour éléments neutres respectifs 0 et 1.
2. La loi de composition interne '-' sur $\mathbb{R}$ n'est ni associative, ni commutative sur $\mathbb{R}$. De plus, elle ne possède pas d'élément neutre.



Si l'élément neutre existe alors il est unique.

Définition 7.3

Un élément $x \in E$ est **inversible** (ou symétrisable) dans E s'il existe au moins un élément $x' \in E$ (dit **inverse** ou **symétrique** de x) tel que $x * x' = x' * x = e$. Nous notons x^{-1} cet inverse.

Exemples

1. Dans $\mathbb{R}$ muni de la loi de composition interne '+', l'inverse de tout élément $x \in \mathbb{R}$ est noté $-x$.
2. Dans $\mathbb{R}^*$ muni de la loi de composition interne '×', l'inverse de tout élément $x \in \mathbb{R}^*$ est noté $\frac{1}{x}$.

Définition 7.4

Si $*$ et T sont deux lois de composition interne de E , $*$ est **distributive** par rapport à T si pour tout $(x; y; z) \in E \times E \times E$,

$$x * (yTz) = (x * y)T(x * z) \text{ et } (yTz) * x = (y * x)T(z * x).$$

2. Groupes**Définition 7.5**

Un ensemble non vide G muni d'une loi de composition interne $*$ est un **groupe** si :

- la loi $*$ est associative,
- il existe un élément neutre e pour la loi $*$,
- tout élément de G possède un symétrique dans G .

Nous notons $(G, *)$ un tel ensemble.

Si de plus la loi est commutative, le groupe est dit **commutatif** ou **abélien**.

Exemples

1. $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$, $(\mathbb{C}, +)$ et $(\mathbb{R}^*, +)$ sont des groupes abéliens.
2. $(\mathbb{N}, +)$ n'est pas un groupe.
3. L'ensemble des homothéties et des translations du plan, muni de la loi de composition interne $\circ$ est un groupe.



Dans un groupe $(G, *)$, tout élément est **régulier** (ou **simplifiable**) c'est-à-dire que nous avons toujours :

$$\forall x, y \in G, x * y = x * z \implies y = z \text{ et } y * x = z * x \implies y = z.$$

Définition 7.6

Soient $(G, *)$ un groupe et H un sous-ensemble de G . H est un **sous-groupe** de $(G, *)$ si

- $\forall (x; y) \in H^2, x * y \in H$,
- $e \in H$,
- $\forall x \in H, x^{-1} \in H$,

avec e l'élément neutre de G et x^{-1} le symétrique de x dans G .

Théorème 7.1

Soient $(G, *)$ un groupe et $((G_i, *))_{i \in I}$ une famille de sous-groupes de $(G, *)$. Alors $(\bigcap_{i \in I} G_i, *)$ est un sous-groupe de $(G, *)$.

3. Morphismes de groupes**Définition 7.7**

Soient $(G, *)$ et (G', T) deux groupes. Une application f de G dans G' est un **morphisme** de groupes si et seulement si $\forall x, y \in G, f(x * y) = f(x)Tf(y)$. Si de plus f est bijective, f est un **isomorphisme** de groupes. Les deux groupes sont alors **isomorphes**.

Exemples

1. $\ln :]0; +\infty[, \times) \rightarrow (\mathbb{R}, +)$ est un morphisme de groupes, c'est même un isomorphisme de groupes. Son application inverse est $\exp$.
2. L'application $x \mapsto 3 \ln x$ réalise un isomorphisme de $(\mathbb{R}_+^*, \times)$ sur $(\mathbb{R}, +)$.

Théorème 7.2

Soient $(G_1, T_1), (G_2, T_2)$ et (G_3, T_3) trois groupes et $f : G_1 \rightarrow G_2$ et $g : G_2 \rightarrow G_3$ deux morphismes de groupes. Alors $g \circ f$ est un morphisme de groupes. Si f et g sont des isomorphismes de groupes, alors $f \circ g$ l'est aussi.

Théorème 7.3

Soient $(G, *)$ et (G', T) deux groupes d'éléments neutres respectifs e et e' , et f un morphisme de G dans G' . Nous avons :

$$e' = f(e) \text{ et } \forall x \in G, f(x^{-1}) = (f(x))^{-1}.$$

- $f(G)$ est un sous-groupe de G' appelé **image** de f et noté $\text{Im} f$.
- L'ensemble $N = f^{-1}(\{e'\}) = \{x \in G; f(x) = e'\}$ est un sous-groupe de G que nous appelons le **noyau** du morphisme f . Nous le notons $\text{Ker} f$.



L'application f est injective si et seulement si $\text{Ker} f = \{e\}$.

Application

D'après le concours d'inspecteur des impôts, 2009

Soit $(G, \cdot)$ un groupe d'élément neutre e . Nous disons que N est un sous-groupe distingué de G si et seulement si

$$\begin{aligned} (sg_1) \quad & \forall (h; h') \in N^2 \quad h \cdot h' \in N \\ (sg_2) \quad & \forall h \in N \quad h^{-1} \in N \\ (sg_3) \quad & \forall (g; h) \in G \times N \quad ghg^{-1} \in N \end{aligned}$$

1. Montrer que G et $\{e\}$ sont des sous-groupes distingués de G .
2. Démontrer que si G est un groupe commutatif alors tout sous-groupe est un sous-groupe distingué.
3. Soient H un sous-groupe de G et N un sous-groupe distingué de G .
 - a. Démontrer que $N \cap H$ est un sous-groupe distingué de H .
 - b. En déduire que si $N \subset H$ alors N est un sous-groupe distingué de H .
4. Démontrer que l'intersection de sous-groupes distingués de G est un sous-groupe distingué de G .

1. Soit $(h \cdot h') \in G^2, hh' \in G, (sg_1)$. Soit $h \in G, h^{-1} \in G, (sg_2)$. Soit $(gh) \in G \times G, ghg^{-1} \in G, (sg_3)$. $e \cdot e = e \in \{e\}, (sg_1)$. $e \cdot e^{-1} \in \{e\}, (sg_2)$. $\forall g \in G, geg^{-1} = e \in \{e\}, (sg_3)$.

2. Soit N un sous-groupe de G . $\forall (hh') \in N^2, \underbrace{hh'}_{N \text{ sous-groupe de } G}, (sg_1)$. $\forall h \in N,$

$$\underbrace{h^{-1}}_{N \text{ sous-groupe de } G} \in N, (sg_2). \forall g \in G, \forall h \in N, \underbrace{ghg^{-1} = gg^{-1}h = h}_{G \text{ est commutatif}} \in N, (sg_3).$$

3. Soient H un sous-groupe de G et N un sous-groupe distingué de G .

a. En premier lieu, $(N \cap H) \subset H$.

- $\forall (a; b) \in (N \cap H)^2$, $ab \in N$ car $a \in N$ et $b \in N$; $\forall (a; b) \in (N \cap H)^2$, $ab \in H$ car $a \in H$ et $b \in H$; donc $ab \in N \cap H$, (sg_1) .
- $\forall a \in N \cap H$, $a^{-1} \in N$ car $a \in N$; $\forall a \in N \cap H$, $a^{-1} \in H$ car $a \in H$; donc $a^{-1} \in N \cap H$, (sg_2) .
- $\forall g \in H$, $\forall h \in N \cap H$, $ghg^{-1} \in H$ car $g \in H$ et $h \in H$; $h \in N$ et $g \in G$ donc $ghg^{-1} \in N$; donc $ghg^{-1} \in N \cap H$, (sg_3) .

b. Si $N \subset H$, alors $N \cap H = N$. La question **3.a.** permet d'affirmer que $N \cap H$ est un sous-groupe distingué de H donc N est un sous-groupe distingué de H .

4. Soit $(IM_i)_{i \in I}$ une famille de sous-groupes distingués de G .

- Soit $(h; h') \in (\bigcap_{i \in I} G_i)^2$, $\forall i \in I$, $(h, h') \in G_i^2$ donc $\underbrace{hh'}_{(sg_1) \text{ pour } G_i} \in G_i$ ainsi $hh' \in \bigcap_{i \in I} G_i$.
 - Soit $h \in \bigcap_{i \in I} G_i$, $\forall i \in I$, $h \in G_i$ donc $\underbrace{h^{-1}}_{(sg_2) \text{ pour } G_i} \in G_i$ ainsi $h^{-1} \in \bigcap_{i \in I} G_i$.
 - Soit $(g; h) \in G \times (\bigcap_{i \in I} G_i)$, $\forall i \in I$, $\underbrace{ghg^{-1}}_{(sg_3) \text{ pour } G_i} \in G_i$ ainsi $ghg^{-1} \in \bigcap_{i \in I} G_i$.
- $\bigcap_{i \in I} G_i$ vérifie (sg_1) , (sg_2) et (sg_3) : $\bigcap_{i \in I} G_i$ est un sous-groupe distingué de G .

Pour s'entraîner (solutions p. 508)

7.1 Sur $\mathbb{R}$, nous définissons la loi $*$ par :

$$\forall a, b \in \mathbb{R}, a * b = a + b + ab.$$

1. Montrer que $*$ est associative.
2. Montrer que $*$ est commutative.
3. Montrer que $*$ possède un élément neutre.

7.2 Nous considérons $G =]-1; 1[$ muni de la loi définie par $x * y = \frac{x+y}{1+xy}$. G est-il un groupe ?

7.3 $\mathbb{Z}^*$ muni de la multiplication usuelle est-il un groupe ?

7.4 Montrer que $\mathbb{N}$ muni de la loi $+$ n'est pas un groupe.

7.5

www

1. Soit $n \in \mathbb{N}$. Résoudre dans $\mathbb{C}$ l'équation suivante : $1 + z + z^2 + \dots + z^n = 0$. Notons G l'ensemble des solutions de cette équation.

2. Rappeler la définition d'un groupe.

3. Montrer que, en adjoignant à G un nombre complexe bien choisi u , $G \cup \{u\}$ est un sous-groupe fini du groupe multiplicatif $\mathbb{C}$.

7.6 Soit $(F, *)$ un groupe. Nous appelons **centre** de F et nous notons $Z(F)$ l'ensemble $\{x \in F; \forall y \in F, x * y = y * x\}$.

1. Montrer que $Z(F)$ est un sous-groupe de F .
2. Montrer que F est commutatif si et seulement si $Z(F) = F$.

Structures algébriques : anneaux et corps

En algèbre, une structure algébrique est une structure formée d'un ensemble associé à une ou plusieurs lois de composition.

1. Anneaux

Définition 8.1

Un ensemble A , muni d'une loi notée $+$ (dite addition) et d'une loi notée $\times$ (dite multiplication) possède une structure d'**anneau** si :

- $(A, +)$ est un groupe abélien. L'élément neutre de $(A, +)$ est noté 0_A .
- la multiplication est associative et possède un élément neutre, noté 1_A .
- la multiplication est distributive par rapport à l'addition.

Nous notons $(A, +, \times)$ cet anneau.

Si la multiplication est commutative, l'anneau est dit **commutatif**.

Exemples

1. $(\mathbb{Z}, +, \times)$ est un anneau pour les lois usuelles $+$ et $\times$.
2. $(\mathbb{Q}, +, \times)$ est un anneau pour les lois usuelles $+$ et $\times$.
3. $(\mathbb{R}, +, \times)$ est un anneau pour les lois usuelles $+$ et $\times$.
4. $(\mathbb{C}, +, \times)$ est un anneau pour les lois usuelles $+$ et $\times$.

Théorème 8.1

Soient $(A, +, \times)$ un anneau et deux éléments x et y qui commutent : $xy = yx$. Pour tout $k \in \mathbb{N}$ et $x \in A$, nous notons $x^k = \underbrace{x \times \dots \times x}_{k \text{ fois}}$, avec la convention $x^0 = 1_A$.

Pour $n \in \mathbb{N}$ et $x, y \in A$, nous avons la **formule du Binôme de Newton** :

$$(x + y)^n = \sum_{k=0}^n \binom{n}{k} x^k y^{n-k} \quad \text{où} \quad \binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

Définition 8.2

Soient $(A, +, \times)$ un anneau et B un sous-ensemble de A . B est un **sous-anneau** de A si

- $(B, +)$ est un sous-groupe abélien de $(A, +)$.
- $1_A \in B$.
- $\forall x, y \in B, x \times y \in B$.

Exemple

$\mathbb{Z}$ est un sous-anneau de $\mathbb{Q}$.

**Point méthode**

Pour montrer qu'une partie non vide B d'un anneau A est un sous-anneau de A , il faut et il suffit que $1_A \in B$ et que :

$$\forall x \in B \quad \forall y \in B, \quad x - y \in B \quad \text{et} \quad x \times y \in B.$$

Définition 8.3

Soient A et B étant deux anneaux. Une application f de A dans B , est un **morphisme d'anneaux** si et seulement si :

$$\forall x, y \in A, f(x + y) = f(x) + f(y) ; f(x \times y) = f(x) \times f(y) ; f(1_A) = 1_B.$$

Définition 8.4

Lorsqu'il existe dans un anneau A des éléments a et b tels que $a \neq 0_A, b \neq 0_A$ et $a \times b = 0_A$, nous disons que a et b sont des **diviseurs de zéro**.

Un anneau **intègre** est un anneau commutatif, non réduit à $\{0_A\}$ et sans diviseur de zéro.

Exemple

$(\mathbb{Z}, +, \times)$ est intègre et ses éléments inversibles sont 1 et -1 .



Pour qu'un anneau $(A, +, \times)$ commutatif, non réduit à $\{0_A\}$, soit intègre, il faut et il suffit que tout élément non nul soit simplifiable pour la multiplication, c'est-à-dire que si a est un élément non nul de A , si b et c sont deux éléments de A tels que $ab = ac$ (resp. $ba = ca$), alors b est égal à c .

2. Corps**Définition 8.5**

L'ensemble $\mathbb{K}$, muni de deux lois de composition interne $+$ et $\times$ est un **corps** si

- $(\mathbb{K}, +, \times)$ est un anneau.
- $(\mathbb{K}^*, \times)$ est un groupe où $\mathbb{K}^* = \mathbb{K} - \{0_{\mathbb{K}}\}$.

Exemples

1. $(\mathbb{Q}, +, \times)$, $(\mathbb{R}, +, \times)$ et $(\mathbb{C}, +, \times)$ sont des corps pour les lois $+$ et $\times$ usuelles.
2. $(\mathbb{Z}, +, \times)$ n'est pas un corps pour les lois $+$ et $\times$ usuelles.
3. L'ensemble des fractions rationnelles constitue un corps.



Dans cet ouvrage, tous les corps seront supposés commutatifs sans avoir besoin de le préciser à chaque fois.



Fiche 13

Définition 8.6

Soient $(\mathbb{K}, +, \times)$ un corps et $\mathbb{L}$ une partie de $\mathbb{K}$. $(\mathbb{L}, +, \times)$ est un **sous-corps** de $\mathbb{K}$ si

- $(\mathbb{L}, +, \times)$ est un sous-anneau de $\mathbb{K}$, non réduit à $\{0_{\mathbb{K}}\}$.
- $\forall x \in \mathbb{L}$ avec $x \neq 0_{\mathbb{K}}$, l'inverse $\frac{1}{x} \in \mathbb{L}$.



Point méthode

Pour montrer qu'une partie non vide $\mathbb{L}$ d'un corps $\mathbb{K}$ est un sous-corps de $\mathbb{K}$, il faut et il suffit que $1_{\mathbb{K}} \in \mathbb{L}$ et que :

$$\begin{aligned} \forall x \in \mathbb{L}, \quad \forall y \in \mathbb{L}, \quad x + y \in \mathbb{L} \quad \text{et} \quad x \times y \in \mathbb{L} \\ \forall x \in \mathbb{L}^*, \quad x^{-1} \in \mathbb{L}^* \quad \text{où} \quad \mathbb{L}^* = \mathbb{L} - \{0\} \end{aligned}$$

Application

Démontrer la formule du binôme de Newton énoncée dans la partie cours :

$$(x + y)^n = \sum_{k=0}^n \binom{n}{k} x^k y^{n-k} \quad \text{où} \quad \binom{n}{k} = \frac{n!}{k!(n-k)!}.$$



Cette démonstration se fait par récurrence et en utilisant le triangle de Pascal.

Fiche 71

- Pour $n = 0$, $(x + y)^0 = 1 = \binom{0}{0} x^0 y^0$. Donc la formule est vraie à l'ordre 0.
- Pour $n = 1$, $\sum_{k=0}^1 \binom{1}{k} x^k y^{1-k} = \binom{1}{0} x^0 y^{1-0} + \binom{1}{1} x^1 y^{1-1} = y + x$.

Donc la formule du binôme est vraie à l'ordre $n = 1$.

- Supposons maintenant qu'elle soit vraie à l'ordre $n \geq 1$ fixé. Montrons qu'elle reste vraie à l'ordre $n + 1$.

$$\begin{aligned} (x + y)^{n+1} &= (x + y) \times (x + y)^n \\ &= (x + y) \times \sum_{k=0}^n \binom{n}{k} x^k y^{n-k} \\ &= \sum_{k=0}^n \binom{n}{k} x^{k+1} y^{n-k} + \sum_{k=0}^n \binom{n}{k} x^k y^{n-k+1} \\ &= \sum_{l=1}^{n+1} \binom{n}{l-1} x^l y^{n+1-l} + \sum_{k=0}^n \binom{n}{k} x^k y^{n+1-k} \\ &= \binom{n}{n} x^{n+1} + \sum_{k=1}^n \left(\binom{n}{k-1} + \binom{n}{k} \right) x^k y^{n+1-k} + \binom{n}{0} y^{n+1} \\ &= \binom{n+1}{n+1} x^{n+1} + \sum_{k=1}^n \binom{n+1}{k} x^k y^{n+1-k} + \binom{n+1}{0} y^{n+1} \\ &= \sum_{k=0}^{n+1} \binom{n+1}{k} x^k y^{n+1-k} \end{aligned}$$



Nous avons utilisé le triangle de Pascal, $\binom{n}{k-1} + \binom{n}{k} = \binom{n+1}{k}$ et le fait que $\binom{n}{n} = \binom{n+1}{n+1}$ et $\binom{n}{0} = \binom{n+1}{0}$.

Donc la formule du binôme est également vraie à l'ordre $n + 1$.

Le raisonnement par récurrence permet donc de conclure que la formule est vraie pour tout $n \in \mathbb{N}$.

Pour s'entraîner (solutions p. 508)

8.1 Soit $(A, +, \times)$ un anneau commutatif. L'élément $x \in A$ est **nilpotent** si et seulement s'il existe $n \in \mathbb{N}^*$ tel que $x^n = 0$.

1. Montrer que, si x est nilpotent, alors $1 - x$ est inversible pour la loi $\times$.

2. Montrer que, si x et y sont nilpotents alors $x \times y$ est nilpotent et $x + y$ aussi.

8.2 Soient $(A, +, \times)$ un anneau intègre et $x \in A$. Montrer que si x est inversible à droite alors x est inversible à gauche pour la loi $\times$.

8.3 Soit $(A, +, \times)$ un anneau. Nous appelons **centre** de A l'ensemble $H = \{x \in A; \forall y \in A, xy = yx\}$. Montrer que H est un sous-anneau de A .

8.4 Soient $(A, +_A, \times_A)$, $(B, +_B, \times_B)$ et $(C, +_C, \times_C)$ des anneaux et $f : A \rightarrow B$ et $g : B \rightarrow C$ deux morphismes d'anneaux. Montrer que la composée $g \circ f$ est un morphisme d'anneaux.

8.5 Montrer que $\mathbb{Z}$ n'est pas un corps pour les lois usuelles.

8.6 

Montrer que le corps des rationnels $\mathbb{Q}$ n'admet pas d'autre sous-corps que lui-même.

8.7 

Soit A l'ensemble des réels suivant :

$$A = \{m + n\sqrt{6}; m, n \in \mathbb{Z}\}.$$

1. Montrer que $(A, +, \times)$ est un sous-anneau de $(\mathbb{R}, +, \times)$.

2. Nous considérons l'application $\varphi : A \rightarrow A$, qui à $m + n\sqrt{6}$ associe $\varphi(m + n\sqrt{6}) = m - n\sqrt{6}$. Montrer que φ est un automorphisme de l'anneau $(A, +, \times)$.

1. Généralités

$\mathbb{Z} = \{\dots; -3; -2; -1; 0; 1; 2; 3; \dots\}$ définit l'**ensemble des entiers relatifs**, muni des lois de composition interne addition $+$ et multiplication $\times$.

Définition 9.1

Si $(a; b) \in \mathbb{Z} \times \mathbb{Z}$, b **divise** a si et seulement s'il existe $q \in \mathbb{Z}$ tel que $a = bq$. Nous disons que a est un **multiple** de b ou que b est un **diviseur** de a . La relation de divisibilité est une relation d'ordre partiel dans $\mathbb{N}$.



La relation de divisibilité n'est pas une relation d'ordre partiel dans $\mathbb{Z}$ car elle n'est pas antisymétrique puisque $1|-1$ et $-1|1$ alors que $1 \neq -1$.

Définition 9.2

Division euclidienne. Pour tout $(a; b) \in \mathbb{Z} \times \mathbb{N}^*$, il existe un couple unique $(q; r) \in \mathbb{Z} \times \mathbb{N}$ tel que :

$$a = bq + r \text{ avec } 0 \leq r < b.$$

q est le **quotient** et r le **reste** de la **division euclidienne** de a par b .

2. PGCD, PPCM

Définition 9.3

Soient a et b deux entiers relatifs non nuls. L'ensemble des nombres de $\mathbb{N}^*$ qui divisent à la fois a et b admet un plus grand élément d . C'est le **plus grand commun diviseur** de a et de b . Nous le notons $PGCD(a; b)$ ou $a \wedge b$.

Théorème 9.1 Algorithme d'Euclide

Soient $(a; b) \in \mathbb{Z} \times \mathbb{N}^*$ et q_1 et r_1 respectivement le quotient et le reste de la division euclidienne de a par b tels que $a = bq_1 + r_1$ avec $0 \leq r_1 < b$. Nous avons :

$$a \wedge b = b \wedge r_1.$$

Nous recommençons ce raisonnement avec b et r_1 . Le dernier reste non nul de ce processus est le **PGCD** de a et de b .

Définition 9.4

- Si $\text{PGCD}(a; b) = 1$, nous disons que a et b sont **premiers** entre eux.
- Soient a et b deux entiers relatifs non nuls et $r = \frac{a}{b}$ un nombre rationnel. Si $d \in \mathbb{N}^*$ désigne le PGCD de a et b , nous avons $a = da'$ et $b = db'$, avec a' et b' premiers entre eux. Nous pouvons alors écrire $r = \frac{a'}{b'}$ (avec $b' \neq 0$). C'est la **forme irréductible** de r .

Définition 9.5

Soient a et b deux entiers relatifs non nuls. L'ensemble des nombres de $\mathbb{N}^*$ qui sont multiples à la fois de a et de b admet un plus petit élément m . C'est le **plus petit commun multiple** de a et de b . Nous le notons $\text{PPCM}(a; b)$ ou $a \vee b$.

Théorème 9.2

Soient a et b deux entiers relatifs non nuls. Alors, nous avons :

$$\text{PGCD}(a; b) \times \text{PPCM}(a; b) = |ab|.$$

Théorème 9.3 Théorème de Bézout

Pour que deux entiers relatifs non nuls a et b soient premiers entre eux il faut et il suffit qu'il existe u et v dans $\mathbb{Z}$ tels que :

$$au + bv = 1.$$

Théorème 9.4 Théorème de Gauss

Soient a, b, c trois entiers relatifs non nuls tels que a divise bc et a est premier avec b , alors a divise c .

3. Anneau $\mathbb{Z}/n\mathbb{Z}$ **Définition 9.6**

Soient $n \in \mathbb{N}^*$ et a et b deux entiers relatifs. L'entier relatif a est **congru à b modulo n** que nous notons $a \equiv b[n]$ si et seulement si n divise $(b - a)$.

Nous désignons par $\mathbb{Z}/n\mathbb{Z}$ l'ensemble des classes ainsi formées par regroupement :

$$\bar{a} = \{b \in \mathbb{Z}; a \equiv b[n]\}.$$



La relation $a \equiv b[n]$ est une relation d'équivalence.

Théorème 9.5

- Pour $n \geq 2$, l'ensemble $\mathbb{Z}/n\mathbb{Z}$ muni des deux lois :

$$\bar{a} + \bar{b} = \overline{a + b} ; \bar{a} \times \bar{b} = \overline{a \times b},$$

est un anneau commutatif.

- Un élément $\bar{a}$ de $\mathbb{Z}/n\mathbb{Z}$ est inversible si et seulement si a et n sont premiers entre eux.



La vérification du fait que les lois $+$ et $\times$ sont bien définies sur $\mathbb{Z}/n\mathbb{Z}$ n'est pas évidente.

Proposition 9.1

$\mathbb{Z}/n\mathbb{Z}$ est un corps si et seulement si n est premier.

Application

1. Déterminer, de deux manières possibles, le *PGCD* de 15 et 12.
2. Soit $n \in \mathbb{N}^*$. Montrer que la relation définie sur $\mathbb{Z}$ par $a \equiv b[n] \iff b - a \in n\mathbb{Z}$ est une relation d'équivalence.

1.

• Méthode 1

Les diviseurs positifs de 15 sont : 1, 3, 5 et 15.

Les diviseurs positifs de 12 sont : 1, 2, 3, 4, 6 et 12.

Les diviseurs positifs communs à 15 et à 12 sont : 1 et 3. Nous en déduisons :

$$\text{PGCD}(15; 12) = 3.$$



Cette méthode n'est cependant pas envisageable lorsque nous cherchons le *PGCD* de grands nombres.

• Méthode 2

En utilisant l'algorithme d'Euclide. Nous avons, en effectuant la division euclidienne de 15 par 12, puis de 12 par 3 :

$$15 = 12 \times 1 + 3,$$

$$12 = 3 \times 4 + 0,$$

avec 1 et 3 respectivement le quotient et le reste de la première division, 4 et 0 le quotient et le reste de la deuxième division.



L'algorithme précise que nous devons choisir le dernier reste avant 0.

Nous en déduisons que :

$$\text{PGCD}(15; 12) = 3.$$

2. Tout d'abord, par définition, nous avons $a \equiv b[n]$ si et seulement si $a - b$ est divisible par n .

Ensuite, nous montrons facilement que la relation est :

- réflexive : pour tout $a \in \mathbb{Z}$, n divise $a - a = 0$, donc $a \equiv b[n]$.
- symétrique : pour tout $a, b \in \mathbb{Z}$ tels que $a \equiv b[n]$, n divise $a - b$ donc n divise $b - a$ et $b \equiv a[n]$.
- transitive : pour tout $a, b, c \in \mathbb{Z}$ tels que $a \equiv b[n]$ et $b \equiv c[n]$, alors n divise $a - b$ et $b - c$ donc n divise $a - c = (a - b) + (b - c)$, et finalement $a \equiv c[n]$.

La relation étant bien réflexive, symétrique et transitive, nous en déduisons que c'est une relation d'équivalence.

Pour s'entraîner (solutions p. 509)

9.1 En utilisant l'algorithme d'Euclide, déterminer les *PGCD* de

1. 853 et 212,
2. 385 et 330,
3. 1395 et 1054.

9.2 Déterminer les triplets $(a; b; c) \in (\mathbb{N}^*)^3$ qui vérifient simultanément

1. $PPCM(a; b) = 42$
2. $PGCD(a; c) = 3$
3. $a + b + c = 29$.

9.3 Déterminer le *PGCD* de $12n^2 + 16n + 5$ et $6n + 5$.

9.4 Montrer que tout nombre pair vérifie $x^2 \equiv 0[8]$ ou $x^2 \equiv 4[8]$.

9.5 Résoudre dans $\mathbb{Z}$ l'équation : $1665x + 1035y = 45$.

9.6



Déterminer tous les couples $(u; v)$ d'entiers relatifs tels que :

$$143u + 100v = 1.$$

9.7



Dans $\mathbb{Z}/60\mathbb{Z}$, montrer que $\overline{11}$ et $\overline{7}$ sont inversibles et déterminer leurs inverses.

9.8



Soient a et b deux entiers naturels non nuls. Notons d leur *PGCD* et m leur *PPCM*.

Déterminer tous les couples $(a; b)$ vérifiant le système :

$$\begin{cases} m = d^2 \\ m + d = 156 \\ a \geq b. \end{cases}$$

Tous les vecteurs considérés dans cette fiche sont des éléments de $\mathbb{R}^n$.

1. Généralités

Définition 10.1

D'un point de vue formel, un **vecteur** est une classe d'équivalence de bipoints pour la relation d'équipolence, c'est-à-dire de segments orientés sur lesquels s'ajoute l'information sur le sens.

D'un point de vue cartésien, c'est un objet défini par ses coordonnées dans un repère. En physique, c'est un objet géométrique fixe défini par sa direction, son sens et sa longueur.

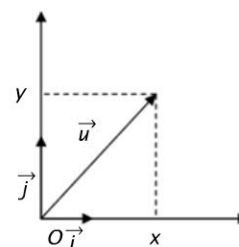
Définition 10.2 Base et coordonnées

Considérons n vecteurs de directions différentes, $\vec{u}_1, \dots, \vec{u}_n$, tels qu'un vecteur $\vec{u}$ quelconque de $\mathbb{R}^n$ puisse s'exprimer de manière unique comme une combinaison linéaire des vecteurs $\vec{u}_1, \dots, \vec{u}_n$, c'est-à-dire qu'il existe une unique suite de scalaires $\lambda_1, \dots, \lambda_n$ tels que $\vec{u} = \lambda_1 \vec{u}_1 + \dots + \lambda_n \vec{u}_n$.

Nous disons que $(O; \vec{u}_1; \dots; \vec{u}_n)$ est un **repère** de $\mathbb{R}^n$, $\{\vec{u}_1; \dots; \vec{u}_n\}$ représente la **base** de ce repère et $\lambda_1, \dots, \lambda_n$ sont les **coordonnées** ou **composantes** du vecteur $\vec{u}$ dans cette base.

Exemple

Dans le repère $(O; \vec{i}; \vec{j})$ du plan, $\vec{u} = x\vec{i} + y\vec{j}$, a pour coordonnées ou composantes x et y .



2. Produit scalaire, produit vectoriel et produit mixte

Définition 10.3

Le **produit scalaire** de deux vecteurs $\vec{u} \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$ et $\vec{v} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$ de $\mathbb{R}^n$, noté par $\vec{u} \cdot \vec{v}$, ou encore $\langle \vec{u} | \vec{v} \rangle$ est le réel

$$\vec{u} \cdot \vec{v} = u_1 v_1 + \dots + u_n v_n = \sum_{i=1}^n u_i v_i,$$

où les u_i (resp. les v_i) pour $i \in \{1, \dots, n\}$ sont les composantes de $\vec{u}$ (resp. $\vec{v}$) dans $\mathbb{R}^n$.

Théorème 10.1

Soient $\vec{u}$ et $\vec{v}$ deux vecteurs de $\mathbb{R}^n$.

- $\vec{u} \cdot \vec{0} = \vec{0} \cdot \vec{u} = 0$.
- $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$ (symétrie du produit scalaire).
- $\vec{u} \cdot \vec{u} = \sum_{i=1}^n u_i^2 \geq 0$.
- Le produit scalaire permet de calculer la **longueur** ou **norme** d'un vecteur. En effet, la longueur d'un vecteur $\vec{u}$ de $\mathbb{R}^n$ est égale à

$$\|\vec{u}\| = \sqrt{\vec{u} \cdot \vec{u}}.$$

- **Inégalité de Cauchy-Schwarz.** Pour tous vecteurs $\vec{u}$ et $\vec{v}$ de $\mathbb{R}^n$, nous avons :

$$|\vec{u} \cdot \vec{v}| \leq \|\vec{u}\| \times \|\vec{v}\|.$$

Définition 10.4

- Deux vecteurs $\vec{u}$ et $\vec{v}$ de $\mathbb{R}^n$ sont **colinéaires** s'il existe $\lambda \in \mathbb{R}$ tel que $\vec{u} = \lambda \vec{v}$ ou $\vec{v} = \lambda \vec{u}$.
- Deux vecteurs non colinéaires $\vec{u}$ et $\vec{v}$ de $\mathbb{R}^n$ engendrent un **plan**. Leur produit scalaire peut être donné en fonction de l'angle orienté $(\vec{u}, \vec{v})$ et de la longueur de chaque vecteur. Si θ est une mesure en radians de l'angle orienté $(\vec{u}, \vec{v})$, alors $\vec{u} \cdot \vec{v} = \|\vec{u}\| \times \|\vec{v}\| \cos \theta$.
- Deux vecteurs $\vec{u}$ et $\vec{v}$ de $\mathbb{R}^n$ sont dit **orthogonaux** si $\vec{u} \cdot \vec{v} = 0$.
- Un **vecteur unitaire** de $\mathbb{R}^n$ est un vecteur dont la norme vaut 1.
- Dans $\mathbb{R}^n$, la base $\{\vec{u}_1; \dots; \vec{u}_n\}$ est dite **orthonormée** si $\forall i \in \{1, \dots, n\}, \|\vec{u}_i\| = 1$ et $\forall i \in \{1, \dots, n\}, j = 1; \dots; n, \vec{u}_i \cdot \vec{u}_j = 0 \quad \forall i \neq j$.

Exemple : Travail d'une force

Soit $\vec{F}$, une force uniforme (c'est-à-dire ayant la même valeur en tous les points de l'espace), d'intensité F et dont le point d'application se déplace suivant le vecteur $\vec{L}$ de longueur L , dans une direction faisant un angle ϕ avec la direction de la force. Alors la force effectue un travail défini par $\vec{F} \cdot \vec{L} = FL \cos \phi$.

Pour définir le produit vectoriel et le produit mixte, nous nous plaçons dans $\mathbb{R}^3$ muni d'une base orthonormée directe.

Définition 10.5

Soient $\vec{u}$ et $\vec{v}$ deux vecteurs de $\mathbb{R}^3$ non colinéaires. Le **produit vectoriel** $\vec{u}$ par $\vec{v}$ est l'unique vecteur $\vec{w}$ de $\mathbb{R}^3$, noté $\vec{u} \wedge \vec{v}$, orthogonal à $\vec{u}$ et à $\vec{v}$ tel que la base $(\vec{u}; \vec{v}; \vec{u} \wedge \vec{v})$ soit directe (c'est-à-dire vérifiant la règle du bonhomme d'Ampère) et de norme $\|\vec{u} \wedge \vec{v}\| = \|\vec{u}\| \times \|\vec{v}\| \times |\sin \theta|$ avec θ l'angle entre $\vec{u}$ et $\vec{v}$.

Exemple : Force de Laplace

La force de Laplace est une force électromagnétique qui s'exerce sur l'ensemble des charges d'un matériau conducteur. Elle résulte de l'action de la force de Lorentz sur toutes les particules chargées. Les matériaux conducteurs étant neutres électriquement, son expression ne dépend que du vecteur $\vec{l}$ dans lequel passe le courant électrique I et du champ magnétique $\vec{B}$. Son expression est $d\vec{F} = I \times d\vec{l} \wedge \vec{B}$.



- Le produit vectoriel est antisymétrique. Pour tous vecteurs $\vec{u}$ et $\vec{v}$ de $\mathbb{R}^3$, $\vec{v} \wedge \vec{u} = -\vec{u} \wedge \vec{v}$.
- Si $\vec{u}$ et $\vec{v}$ sont colinéaires, alors le produit vectoriel de $\vec{u}$ par $\vec{v}$ est le vecteur nul de $\mathbb{R}^3$, $\vec{u} \wedge \vec{v} = \vec{0}$.
- Pour déterminer le sens du produit vectoriel $\vec{u} \wedge \vec{v}$, nous pouvons utiliser la règle de la main droite. Le pouce indiquant le sens de $\vec{u}$ et l'index indiquant le sens de $\vec{v}$, le majeur indique alors le sens de $\vec{u} \wedge \vec{v}$.

Définition 10.6

Le **produit mixte** de $\vec{u}$, $\vec{v}$ et $\vec{w}$, trois vecteurs de $\mathbb{R}^3$, noté par $(\vec{u}, \vec{v}, \vec{w})$ est le produit scalaire de $\vec{w}$ par le produit vectoriel $(\vec{u} \wedge \vec{v})$:

$$(\vec{u}, \vec{v}, \vec{w}) = (\vec{u} \wedge \vec{v}) \cdot \vec{w}.$$



Soient $\vec{u}$, $\vec{v}$ et $\vec{w}$ trois vecteurs de $\mathbb{R}^3$, alors

1. échanger deux vecteurs change le signe du produit mixte :

$$(\vec{u}, \vec{v}, \vec{w}) = -(\vec{v}, \vec{u}, \vec{w}) = -(\vec{w}, \vec{v}, \vec{u}) = -(\vec{u}, \vec{w}, \vec{v}).$$

2. une permutation circulaire des trois vecteurs laisse inchangé le produit mixte :

$$(\vec{u}, \vec{v}, \vec{w}) = (\vec{v}, \vec{w}, \vec{u}) = (\vec{w}, \vec{u}, \vec{v}).$$

Exemple : Force de Lorentz

La relation de Lorentz exprime la force magnétique $\vec{F}$ exercée sur une particule de charge électrique q , animée d'une vitesse $\vec{v}$ dans un champ magnétique $\vec{B}$: $\vec{F} = q\vec{v} \wedge \vec{B}$. La puissance P de la force de Lorentz, définie par $\vec{F} \cdot \vec{v}$ est nulle car la force $\vec{F}$ est constamment perpendiculaire au vecteur vitesse $\vec{v}$ de la particule.

$$P = \vec{F} \cdot \vec{v} = (q\vec{v} \wedge \vec{B}) \cdot \vec{v} = q(\vec{v}, \vec{B}, \vec{v}) = 0.$$

Soient $(O, \vec{i}, \vec{j}, \vec{k})$ le repère orthonormé de $\mathbb{R}^3$ et $\vec{u} = \vec{i} + 4\vec{j} - 2\vec{k}$ et $\vec{v} = -\vec{i} - 3\vec{j} + \vec{k}$ deux vecteurs. Le couple $(\vec{u}, \vec{v})$ définit un plan car $\vec{u}$ et $\vec{v}$ ne sont pas colinéaires. Trouver un vecteur $\vec{w}$ orthogonal au plan défini par $(\vec{u}, \vec{v})$.



Par définition, le produit vectoriel $\vec{u} \wedge \vec{v}$ est orthogonal à $\vec{u}$ et à $\vec{v}$ et donc au plan qu'ils définissent.

Nous choisissons $\vec{w}$ comme tout vecteur colinéaire à $\vec{u} \wedge \vec{v}$.

$$\text{De plus } \vec{u} \wedge \vec{v} = \begin{pmatrix} 1 \\ 4 \\ -2 \end{pmatrix} \wedge \begin{pmatrix} -1 \\ -3 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}.$$

Nous prenons donc $\vec{w} = \lambda \times (-2\vec{i} + \vec{j} + \vec{k})$ avec $\lambda \in \mathbb{R}$.

Le plan $(\vec{u}, \vec{v})$ est défini comme tous les vecteurs normaux à $\vec{u} \wedge \vec{v}$. Si nous posons $\vec{d} = a\vec{i} + b\vec{j} + c\vec{k}$ alors

$$(\vec{u} \wedge \vec{v}) \cdot \vec{d} = 0 \iff -2a + b + c = 0.$$

De plus, $\vec{d}$ est orthogonal à $\vec{u}$ donc

$$\vec{d} \cdot \vec{u} = 0 \iff a + 4b - 2c = 0.$$



Nous avons donc un système de deux équations et trois inconnues à résoudre. Une solution est donnée par $a = 2, b = 1$ et $c = 3$ donc $\vec{d} = 2\vec{i} + \vec{j} + 3\vec{k}$ ou tout autre vecteur proportionnel à celui-ci permet de répondre à la question.

Pour s'entraîner (solutions p. 509)

10.1 Soient A, B et C trois points distincts du plan. Répondre par vrai ou faux aux assertions suivantes.

1. Les points A, B et C sont alignés si et seulement si : $\vec{AB} \cdot \vec{AC} = AB \times AC$.
2. Les droites (AB) et (AC) sont orthogonales si et seulement si $\vec{AB} \cdot \vec{AC} = 0$.
3. A est le milieu du segment $[BC]$ si et seulement si : $\vec{AB} \cdot \vec{AC} = -AB^2$.

10.2 Soient $\vec{u}$ et $\vec{v}$ deux vecteurs de même norme. Démontrer que les vecteurs $\vec{u} + \vec{v}$ et $\vec{u} - \vec{v}$ sont deux vecteurs orthogonaux.

10.3 Soit $(O; \vec{i}; \vec{j}; \vec{k})$ un repère orthonormé de $\mathbb{R}^3$. Nous posons $\vec{u} = \vec{i}, \vec{v} = \vec{i} + \vec{j}$ et $\vec{w} = \vec{j}$. Calculer $(\vec{u} \wedge \vec{v}) \wedge \vec{w}$, puis $\vec{u} \wedge (\vec{v} \wedge \vec{w})$ et conclure.

10.4 Soient $\vec{u}, \vec{v}$ et $\vec{w}$ trois vecteurs de $\mathbb{R}^3$. Montrer l'identité de Jacobi :

$$\vec{u} \wedge (\vec{v} \wedge \vec{w}) + \vec{v} \wedge (\vec{w} \wedge \vec{u}) + \vec{w} \wedge (\vec{u} \wedge \vec{v}) = \vec{0}.$$

10.5 Soient $\vec{u} \begin{pmatrix} 3 \\ -1 \\ 3 \end{pmatrix}, \vec{v} \begin{pmatrix} 2 \\ 5 \\ -4 \end{pmatrix}$ et $\vec{w} \begin{pmatrix} -2 \\ -3 \\ 6 \end{pmatrix}$ trois vecteurs de $\mathbb{R}^3$. Calculer les produits mixtes suivants :

1. $(\vec{u}, \vec{v}, \vec{w})$
2. $(\vec{u}, \vec{w}, \vec{v})$
3. $(\vec{v}, \vec{w}, \vec{u})$
4. $(\vec{u}, \vec{v}, \vec{u})$

Vecteurs et éléments de géométrie

1. Applications du produit scalaire

Définition 11.1

En géométrie analytique, un **repère** vectoriel **orthonormé** du plan $\mathbb{R}^2$ ou de l'espace $\mathbb{R}^3$ est constitué d'un point origine O et d'une base composée de deux vecteurs orthogonaux et unitaires $\vec{i}, \vec{j}$ dans le plan ou de trois vecteurs orthogonaux deux à deux et unitaires $\vec{i}, \vec{j}, \vec{k}$ dans l'espace. Les **coordonnées** d'un point A du plan ou de l'espace dans ce repère sont alors données respectivement dans le plan et l'espace par les réels x_A, y_A, z_A tels que : $\overrightarrow{OA} = x_A \vec{i} + y_A \vec{j}$ ou $\overrightarrow{OA} = x_A \vec{i} + y_A \vec{j} + z_A \vec{k}$.

Nous notons généralement A et ses coordonnées par $A \begin{pmatrix} x_A \\ y_A \end{pmatrix}$ (dans le plan) ou $A \begin{pmatrix} x_A \\ y_A \\ z_A \end{pmatrix}$ (dans l'espace).

Définition 11.2

Deux points du plan $A \begin{pmatrix} x_A \\ y_A \end{pmatrix}$ et $B \begin{pmatrix} x_B \\ y_B \end{pmatrix}$ ou de l'espace $A \begin{pmatrix} x_A \\ y_A \\ z_A \end{pmatrix}$ et $B \begin{pmatrix} x_B \\ y_B \\ z_B \end{pmatrix}$ définissent un vecteur $\overrightarrow{AB} \begin{pmatrix} x_B - x_A \\ y_B - y_A \end{pmatrix}$ dans le plan et $\overrightarrow{AB} \begin{pmatrix} x_B - x_A \\ y_B - y_A \\ z_B - z_A \end{pmatrix}$ dans l'espace. La **distance** entre A et B est donnée par $AB = \|\overrightarrow{AB}\|$.

Définition 11.3 Équation d'une droite dans le plan $\mathbb{R}^2$

Soient $M_0 \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$ un point de $\mathbb{R}^2$ et $\vec{n} \begin{pmatrix} a \\ b \end{pmatrix}$ un vecteur non nul du plan. La droite (D) du plan $\mathbb{R}^2$ passant par M_0 et normale à $\vec{n}$ est l'ensemble des points $M \begin{pmatrix} x \\ y \end{pmatrix}$ tels que $\overrightarrow{M_0M}$ et $\vec{n}$ soient orthogonaux, c'est-à-dire :

$$\overrightarrow{M_0M} \cdot \vec{n} = 0 \iff a(x - x_0) + b(y - y_0) = 0.$$

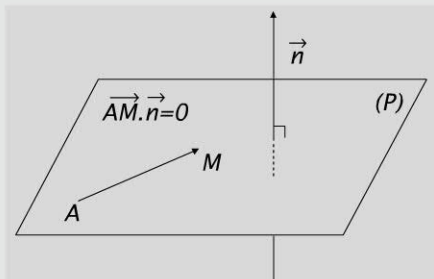
Cette dernière équation est appelée **équation cartésienne** de la droite (D) .

Définition 11.4 Distance d d'un point à une droite dans le plan

Soient a, b deux réels non simultanément nuls. La distance d d'un point $A \begin{pmatrix} x_A \\ y_A \end{pmatrix}$ du plan à la droite d'équation cartésienne $ax + by + c = 0$ est : $d = \frac{|ax_A + by_A + c|}{\sqrt{a^2 + b^2}}$.

Définition 11.5 Équation d'un plan dans l'espace $\mathbb{R}^3$

Soit $M_0 \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix}$ un point de $\mathbb{R}^3$ et $\vec{n} \begin{pmatrix} a \\ b \\ c \end{pmatrix}$ un vecteur non nul de l'espace. Le plan (P) passant par M_0 et de **vecteur normal** $\vec{n}$ est l'ensemble des points $M \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ de l'espace tels que $\overrightarrow{M_0M}$ et $\vec{n}$ soient orthogonaux, c'est-à-dire :



$$\overrightarrow{AM} \cdot \vec{n} = a(x - x_0) + b(y - y_0) + c(z - z_0) = 0.$$

Cette dernière équation est appelée **équation cartésienne** du plan (P) .

Définition 11.6 Distance d d'un point à un plan dans l'espace

Soient a, b, c trois réels non simultanément nuls. La distance d d'un point $A \begin{pmatrix} x_A \\ y_A \\ z_A \end{pmatrix}$ de l'espace au plan d'équation cartésienne $ax + by + cz + d = 0$ est donnée par :

$$d = \frac{|ax_A + by_A + cz_A + d|}{\sqrt{a^2 + b^2 + c^2}}.$$

Définition 11.7 Équations d'une droite dans l'espace $\mathbb{R}^3$

Soient $M_0 \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix}$ un point de $\mathbb{R}^3$ et $\vec{u} \begin{pmatrix} a \\ b \\ c \end{pmatrix}$ un vecteur non nul de l'espace. Chacun des points M appartenant à la droite passant par M_0 et de direction $\vec{u}$ est tel que les vecteurs $\overrightarrow{MM_0}$ et $\vec{u}$ soient colinéaires : $\overrightarrow{MM_0} = \lambda \vec{u}$. Ainsi,

$$M = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} + \lambda \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} x_0 + \lambda a \\ y_0 + \lambda b \\ z_0 + \lambda c \end{pmatrix}$$

Cette dernière équation est appelée **équation paramétrique** de la droite (D) .

Dans l'espace, une droite peut aussi se définir comme l'intersection de deux plans non parallèles :

$$\begin{cases} ax + by + cz + d = 0 \\ a'x + b'y + c'z + d' = 0 \end{cases}$$

Ce dernier système d'équations est appelé **équation cartésienne** de la droite (D) .

2. Applications du produit vectoriel et du produit mixte

Définition 11.8 Distance d d'un point M à une droite (D) dans l'espace

La distance d'un point M à une droite (D) passant par les points distincts A et B est donnée par :

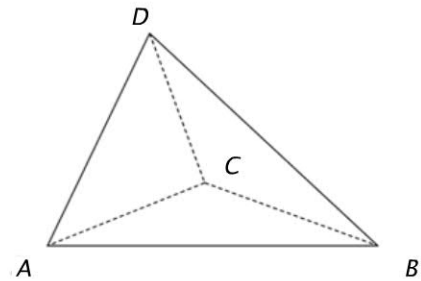
$$d = \frac{\|\vec{MA} \wedge \vec{AB}\|}{\|\vec{AB}\|} = \frac{\|\vec{MA} \wedge \vec{MB}\|}{\|\vec{AB}\|}.$$

Théorème 11.1

- L'aire d'un **parallélogramme** $ABCD$ est $\|\vec{AB} \wedge \vec{AD}\|$.
- L'aire d'un **triangle** ABC est $\frac{1}{2}\|\vec{AB} \wedge \vec{AC}\|$.

Théorème 11.2

- Le volume d'un **parallélépipède** construit sur les vecteurs $\vec{AB}$, $\vec{AD}$ et $\vec{AE}$ est $|(\vec{AB}, \vec{AD}, \vec{AE})|$.
- Le volume d'un **tétraèdre** $ABCD$ est $\frac{1}{6}|(\vec{AB}, \vec{AC}, \vec{AD})|$.

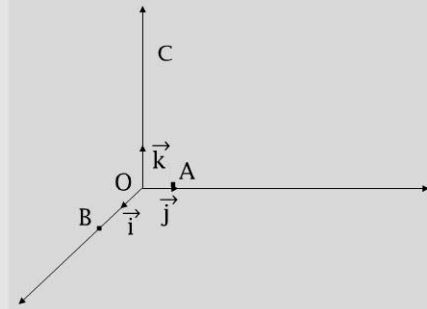


Tétraèdre $ABCD$

Application

Nous travaillons dans le repère $(O, \vec{i}, \vec{j}, \vec{k})$ orthonormé direct. Nous considérons les trois points A , B et C tels que $\vec{OA} = \vec{i}$, $\vec{OB} = 2\vec{j}$ et $\vec{OC} = 3\vec{k}$.

1. Donner l'équation du plan passant par les points A , B et C .
2. Calculer la distance de l'origine à ce plan.
3. Donner l'aire du triangle ABC .



1. Nous avons tout d'abord $\vec{AB} \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix}$ et $\vec{AC} \begin{pmatrix} -1 \\ 0 \\ 3 \end{pmatrix}$. Le vecteur $\vec{n} = \vec{AB} \wedge \vec{AC}$ est orthogonal à ce plan. Calculons ses coordonnées.

$$\vec{n} = \vec{AB} \wedge \vec{AC} = \begin{pmatrix} 6 \\ 3 \\ 2 \end{pmatrix}.$$

Soit maintenant $M(x; y; z)$ qui appartient au plan. Alors, nous avons $\vec{AM} \cdot \vec{n} = 0$, ce qui revient à $6(x - 1) + 3y + 2z = 0$. Nous en déduisons que le plan (P) a pour équation :

$$6x + 3y + 2z - 6 = 0.$$

2. Donnons maintenant la distance de l'origine à ce plan. Elle est donnée par :

$$d = \frac{|ax_0 + by_0 + cz_0 + d|}{\sqrt{a^2 + b^2 + c^2}} = \frac{6}{\sqrt{6^2 + 3^2 + 2^2}} = \frac{6}{7}.$$

3. Nous cherchons pour terminer l'aire du triangle ABC . Elle est donnée par :

$$\mathcal{A} = \frac{1}{2} \|\vec{AB} \wedge \vec{AC}\| = \frac{1}{2} \times \sqrt{36 + 9 + 4} = \frac{7}{2}.$$

Pour s'entraîner (solutions p. 509)

Dans tous les exercices, nous considérons le repère orthonormé direct $(O, \vec{i}, \vec{j}, \vec{k})$ de $\mathbb{R}^3$.

11.1 Déterminer l'angle aigu entre les plans (P) $2x - y + z = 3$ et (P') $x + y + 2z = 7$ dans $\mathbb{R}^3$.

11.2 Donner l'équation du plan de $\mathbb{R}^3$ passant par les points $A \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$, $B \begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix}$ et $C \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}$.

11.3 Donner l'équation du plan passant par le point $A \begin{pmatrix} 2 \\ 1 \\ -3 \end{pmatrix}$ et dont le vecteur normal est $\vec{n} \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$ dans $\mathbb{R}^3$.

11.4 Nous considérons dans $\mathbb{R}^3$ les deux plans $P : x - 4y + 7 = 0$ et $Q : x + 2y - z + 1 = 0$. Les deux plans sont-ils sécants ?

11.5 Nous considérons la sphère S de $\mathbb{R}^3$ de diamètre $[AB]$ avec $A \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}$ et $B \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix}$.

1. Déterminer l'équation de la sphère (S) .
2. Donner le centre et le rayon de cette sphère.

Dans toute cette Fiche, $\mathbb{K}$ désigne un corps commutatif.

1. Définitions

Définition 12.1

Un **polynôme** à une indéterminée, à coefficients dans $\mathbb{K}$, est une suite notée $(a_i)_{i \in \mathbb{N}}$ de scalaires de $\mathbb{K}$ tous nuls à partir d'un certain rang p . Un tel polynôme se note généralement par P ou $P(X)$ et s'écrit :

$$P(X) = a_0 + a_1X + a_2X^2 + \dots + a_pX^p.$$

Les $(a_i)_{i \in \mathbb{N}}$ sont les **coefficients** du polynôme. Deux polynômes seront égaux si et seulement si leurs coefficients sont égaux. Le **degré** du polynôme, noté par $d^\circ(P)$, est la plus grande valeur de p telle que a_p soit non nul. a_p est alors appelé le **coefficient dominant** de P .

Nous convenons également que le polynôme nul a pour degré $-\infty$.

Nous notons par $\mathbb{K}[X]$ l'ensemble des polynômes à une indéterminée à coefficients dans $\mathbb{K}$.

Nous notons également par $\mathbb{K}_n[X]$ l'ensemble des polynômes de degré inférieur ou égal à n .

Définition 12.2

Soient P et Q deux polynômes de $\mathbb{K}[X]$ avec $P(X) = \sum_{i=0}^p a_iX^i$ et $Q(X) = \sum_{j=0}^q b_jX^j$.

- Le **polynôme somme** s'écrit $P+Q$ avec $(P+Q)(X) = \sum_{k=0}^r c_kX^k$ avec $r = \max(p, q)$ et $c_k = a_k + b_k$.
- Soit $\lambda \in \mathbb{K}$. Le **polynôme produit par un scalaire** s'écrit λP avec $(\lambda P)(X) = \sum_{i=0}^p (\lambda a_i)X^i$.
- Le **polynôme produit** s'écrit $PQ \in \mathbb{K}[X]$ avec $(PQ)(X) = \sum_{k=0}^{p+q} d_kX^k$ avec $d_k = \sum_{i+j=k} a_ib_j$.



- $d^\circ(P + Q) \leq \max(d^\circ(P), d^\circ(Q))$.
- Si $\lambda \in \mathbb{K}$ est différent de 0, alors $d^\circ(\lambda P) = d^\circ(P)$, sinon $d^\circ(\lambda P) = d^\circ(0) = -\infty$.
- $d^\circ(PQ) = d^\circ(P) + d^\circ(Q)$.

2. Division euclidienne

Définition 12.3

Soient A et B deux polynômes de $\mathbb{K}[X]$. A divise B dans $\mathbb{K}[X]$ s'il existe un polynôme Q dans $\mathbb{K}[X]$ tel que $B = AQ$.

Théorème 12.1

Soient A et B deux polynômes $\mathbb{K}[X]$ tels que $B \neq 0$. Il existe alors un unique couple (Q, R) de polynômes de $\mathbb{K}[X]$ tel que :

$$A = BQ + R,$$

avec $d^\circ(R) < d^\circ(B)$. Les polynômes Q et R s'appellent respectivement le **quotient** et le **reste** de la **division euclidienne** de A par B .



Deux exemples de calcul de division euclidienne sont détaillés dans l'application.

Définition 12.4

Un polynôme $P \in \mathbb{K}[X]$ est dit **irréductible** dans $\mathbb{K}[X]$ si P n'est pas constant (c'est-à-dire $d^\circ(P) \geq 1$) et si les seuls polynômes divisant P sont les constantes non nulles et les polynômes λP , avec $\lambda \in \mathbb{K}$.

Théorème 12.2

Soit $P \in \mathbb{K}[X]$ un polynôme non nul. Alors P se décompose de manière unique à l'ordre près sous la forme $P = \lambda P_1^{\alpha_1} \dots P_k^{\alpha_k}$ où $\lambda \in \mathbb{K}^*$, $\alpha_i \in \mathbb{N}^*$ et les P_i sont des polynômes distincts, de coefficients dominants égaux à 1 et irréductibles dans $\mathbb{K}[X]$.

Définition 12.5

Deux polynômes P et Q de $\mathbb{K}[X]$ sont **premiers** entre eux s'ils n'ont aucun polynôme commun dans leurs décompositions respectives sous la forme donnée dans le théorème précédent.



La division euclidienne dans $\mathbb{R}[X]$ est l'analogue de la division dans $\mathbb{Z}$. Il y a ainsi des ressemblances fortes entre les notions de nombres premiers, d'irréductibles, de décomposition en produit de nombres premiers ou d'irréductibles en arithmétique et celles introduites ci-dessus.

Fiche 9

3. Racines d'un polynôme

Définition 12.6

Un scalaire a de $\mathbb{K}$ est dit **racine** de P si l'évaluation du polynôme en a est nulle, soit $P(a) = 0$. a est racine de P si et seulement si le polynôme $(X - a)$ divise P . Soit $k \in \mathbb{N}^*$. a est **racine d'ordre k** (ou de **multiplicité k**) de P si le polynôme $(X - a)^k$ divise P et si le polynôme $(X - a)^{k+1}$ ne divise pas P .

Théorème 12.3

Si $a_1, a_2, \dots, a_r$ sont des racines distinctes du polynôme P de $\mathbb{K}[X]$ de multiplicités respectives $k_1, k_2, \dots, k_r$, alors P peut s'écrire sous la forme $P(X) = (X - a_1)^{k_1} \dots (X - a_r)^{k_r} Q(X)$, où Q est un polynôme de $\mathbb{K}[X]$ tel que $Q(a_1) \neq 0, \dots, Q(a_r) \neq 0$.

Corollaire 12.1

Un polynôme de degré n a au plus n racines.

Théorème 12.4

Soient P un polynôme de $\mathbb{K}[X]$ et a un élément de $\mathbb{K}$. Alors, les propriétés suivantes sont équivalentes :

- $(X - a)^k$ divise P .
- $P(a) = P'(a) = \dots = P^{(k-1)}(a) = 0$.

En conséquence, a est une racine d'ordre k de P si et seulement si :

$$P(a) = P'(a) = \dots = P^{(k-1)}(a) = 0, P^{(k)}(a) \neq 0.$$

4. Cas de $\mathbb{C}[X]$ et $\mathbb{R}[X]$

Théorème 12.5 Théorème d'Alembert

Tout polynôme non constant de $\mathbb{C}[X]$ admet au moins une racine dans $\mathbb{C}$.



Tout polynôme de $\mathbb{C}[X]$ de degré n admet exactement n racines (comptées avec leurs multiplicités). Tout polynôme non constant de $\mathbb{C}[X]$ est **scindé**, c'est-à-dire factorisable en un produit de facteurs du premier degré. Nous disons que $\mathbb{C}$ est **algébriquement clos**.

Théorème 12.6

- **Décomposition d'un polynôme.** Soit P un polynôme de $\mathbb{R}[X]$. Nous pouvons considérer P comme un polynôme de $\mathbb{C}[X]$ et à ce titre le décomposer en un produit de facteurs du premier degré de la forme $(X - \alpha_i)$ où les α_i sont les racines complexes de P , comptées avec leur multiplicité.
- Les polynômes **irréductibles** de $\mathbb{R}[X]$ sont soit de la forme $\lambda(X - \alpha)$ où $\lambda, \alpha \in \mathbb{R}$, soit de la forme $\lambda(X^2 - sX + p)$ avec $s^2 - 4p < 0$ où $s, p \in \mathbb{R}$. Tout polynôme de $\mathbb{R}[X]$ peut se décomposer en un produit de tels facteurs irréductibles.

Application

1. Nous considérons les deux polynômes de $\mathbb{R}[X]$, $A(X) = 2X^3 - 11X^2 + 7X + 20$ et $B(X) = 2X - 5$. Chercher les polynômes $Q(X)$ et $R(X)$ de $\mathbb{R}[X]$ tels que :

$$A(X) = B(X)Q(X) + R(X).$$

2. Soient à présent les deux polynômes de $\mathbb{R}[X]$ $A(X) = X^3 + 2X^2 + 3$ et $B(X) = X^2 - X - 1$. Chercher les polynômes $Q(X)$ et $R(X)$ de $\mathbb{R}[X]$ tels que :

$$A(X) = B(X)Q(X) + R(X).$$

1.



Pour cela, nous posons la division de $A(X)$ par $B(X)$ selon les puissances décroissantes de X :

$$\begin{array}{r|rr} 2X^3 & -11X^2 & +7X & +20 & 2X & -5 \\ -2X^3 & +5X^2 & & & X^2 & -3X & -4 \\ \hline & -6X^2 & & & & & \\ & +6X^2 & -15X & & & & \\ \hline & & -8X & & & & \\ & & +8X & -20 & & & \\ \hline & & & 0 & & & \end{array}$$

Nous en déduisons que $Q(X) = X^2 - 3X - 4$ et $R(X) = 0$.

2. En posant à nouveau la division de A par B nous obtenons :

$$\begin{array}{r|l} X^3 + 2X^2 & +3 \\ -X^3 + X^2 + X & X + 3 \\ \hline 3X^2 + X & \\ -3X^2 + 3X + 3 & \\ \hline 4X + 6 & \end{array}$$

Nous en déduisons que $Q(X) = X + 3$ et $R(X) = 4X + 6$.

Pour s'entraîner
(solutions p. 509)

12.1 Donner la division euclidienne du polynôme $A(X)$ par le polynôme $B(X)$ dans $\mathbb{R}[X]$ dans chacun des cas suivants :

1. $A(X) = X^3 + X^2 + X + 1$ et $B(X) = X^2 + 1$.
2. $A(X) = X^3 + 2X^2 + 3$ et $B(X) = X^2 - X - 1$.
3. $A(X) = X^4 + 2$ et $B(X) = X^4 + 4X - 2$.

12.2 Soit le polynôme $P(X) = (X^2 + 1)^2 - (4X + 2)^2$ dans $\mathbb{R}[X]$.

1. Quel est le degré du polynôme $P(X)$?
2. Déterminer les racines réelles de P .

12.3 Soit le polynôme $P(X) = -X^3 + 6X^2 - 9X + k$ dans $\mathbb{R}[X]$ avec k réel.
Pour quelle valeur de k le nombre 4 est-il racine de P ?

12.4 Soit le polynôme $P(X) = X^4 + X^3 - 7X^2 - 13X - 6$ dans $\mathbb{R}[X]$.

1. Quel est le degré du polynôme $P(X)$?
2. Montrer que $X = -1$ est racine de $P(X)$ dans $\mathbb{R}[X]$.
3. Trouver le polynôme $Q(X)$ du troisième degré tel que $P(X) = (X + 1)Q(X)$.
4. Quelles sont les racines réelles de $Q(X)$? En déduire celles de $P(X)$.

12.5 Soit le polynôme $P(X) = 2X^3 - 7X + 2$ dans $\mathbb{R}[X]$.

1. Montrer que -2 est racine de $P(X)$.
2. Factoriser $P(X)$ et donner les racines réelles de P .

Fractions rationnelles

Dans toute cette fiche, $\mathbb{K}$ désigne un corps commutatif.

Définition 13.1

Nous appelons **fraction rationnelle** toute fraction de deux polynômes P et Q appartenant respectivement à $\mathbb{K}[X]$ et $\mathbb{K}^*[X]$, $F(X) = \frac{P(X)}{Q(X)}$. Si P et Q sont premiers entre eux, l'écriture $F = \frac{P}{Q}$ s'appelle la **forme irréductible** de F . L'ensemble des fractions rationnelles est noté $\mathbb{K}(X)$.

Exemple

La fraction $F(X) = \frac{2X+1}{X^3+1}$ est une fraction rationnelle avec $P(X) = 2X+1$ et $Q(X) = X^3+1$.

Définition 13.2

Soit $F = \frac{P}{Q}$ une fraction rationnelle écrite sous forme irréductible. Nous appelons **pôle** de F toute racine de Q . a est un **pôle d'ordre** n de F si a est une racine de multiplicité n de Q . Si $n = 1$, nous disons que a est un **pôle simple** de F .

Théorème 13.1

Soient $F \in \mathbb{K}(X)$, $F \neq 0$ et $F = \frac{P}{Q}$ sa forme irréductible. Soit $Q = \lambda Q_1^{\alpha_1} \dots Q_k^{\alpha_k}$ la décomposition de Q en polynômes irréductibles de $\mathbb{K}[X]$. La **décomposition en éléments simples** de F est l'unique écriture de F sous la forme :

$$F = E + \sum_{i=1}^k \left(\sum_{j=1}^{\alpha_i} \frac{A_{i,j}}{Q_i^j} \right)$$

avec $E \in \mathbb{K}[X]$, $A_{i,j} \in \mathbb{K}[X]$ et $d^\circ(A_{i,j}) < d^\circ(Q_i)$. Le polynôme E s'appelle la **partie entière** de F , il est le quotient de la division euclidienne de P par Q . Les termes $\frac{A_{i,j}}{Q_i^j}$ s'appellent les **éléments simples** de F . Les termes $\sum_{j=1}^{\alpha_i} \frac{A_{i,j}}{Q_i^j}$ s'appellent **partie principale** de F , engendrée par $Q_i^{\alpha_i}$.

1. Décomposition en éléments simples dans $\mathbb{C}(X)$

Nous allons commencer par expliquer comment donner la décomposition d'une fraction rationnelle à coefficients réels en éléments simples sur $\mathbb{C}$.

Théorème 13.2

Soient $F \in \mathbb{C}(X)$, $F \neq 0$ et $F = \frac{P}{Q}$ sa forme irréductible. Toutes les parties principales de F sont de première espèce et F s'écrit de manière unique :

$$F(X) = E(X) + \sum_{i=1}^k \left(\sum_{j=1}^{\alpha_i} \frac{a_{i,j}}{(X - \mu_i)^j} \right)$$

avec $E \in \mathbb{C}[X]$, $a_{i,j} \in \mathbb{C}$ et $\mu_1, \dots, \mu_k$ désignent les pôles de F , d'ordre respectif $\alpha_1, \dots, \alpha_k$.

2. Décomposition en éléments simples dans $\mathbb{R}(X)$

Nous allons maintenant nous intéresser à la décomposition d'une fraction rationnelle à coefficients réels sur $\mathbb{R}$.

Théorème 13.3

Soit $F \in \mathbb{R}(X)$. Les pôles de F sont réels et/ou des couples de complexes conjugués.

- Chaque pôle réel μ , d'ordre n , engendre une partie principale de la forme

$$\frac{a_n}{(X - \mu)^n} + \frac{a_{n-1}}{(X - \mu)^{n-1}} + \dots + \frac{a_2}{(X - \mu)^2} + \frac{a_1}{(X - \mu)},$$

avec $a_1, a_2, \dots, a_n$ des coefficients réels. Ces derniers sont déterminés soit par identification soit en remplaçant X par des valeurs particulières. Ces éléments de décomposition s'appellent les **éléments simples de première espèce**.

- Chaque couple de pôles complexes conjugués $\mu + i\gamma$ et $\mu - i\gamma$, d'ordre n , engendre une partie principale qui s'écrit sous la forme suivante :

$$\frac{a_n X + b_n}{[(X - \mu)^2 + \gamma^2]^n} + \frac{a_{n-1} X + b_{n-1}}{[(X - \mu)^2 + \gamma^2]^{n-1}} + \dots + \frac{a_2 X + b_2}{[(X - \mu)^2 + \gamma^2]^2} + \frac{a_1 X + b_1}{[(X - \mu)^2 + \gamma^2]^1},$$

avec $a_1, a_2, \dots, b_n, b_1, \dots, b_n$ des coefficients réels à déterminer. Ces éléments de décomposition s'appellent les **éléments simples de deuxième espèce**.



Dans le cas où le pôle a est un pôle simple, sa partie principale ne comprend que le terme $\frac{a_1}{(X - a)}$. Le coefficient a_1 est alors déterminé en calculant $(X - a)F(a)$ au point $X = a$.



Point méthode

Pour décomposer une fraction rationnelle $F = \frac{P}{Q}$ en éléments simples, nous pouvons suivre les étapes suivantes :

- recherche de la partie entière (si $d^\circ(F) \geq 0$) ;
- recherche des pôles de F et factorisation de Q ;
- écriture de la forme générale de la décomposition en respectant la nature et l'ordre de multiplicité des pôles ;
- évaluation des divers coefficients (par identification ou toute autre méthode).

Exemples

Voici deux exemples, non évidents, de décompositions en éléments simples dans $\mathbb{R}(X)$:

$$\frac{1}{1+X^4} = \frac{1}{4} \frac{\sqrt{2}X+2}{X^2+\sqrt{2}X+1} + \frac{1}{4} \frac{-\sqrt{2}X+2}{X^2-\sqrt{2}X+1}.$$
$$\frac{X+1}{(X-1)^4(X-2)^2} = \frac{2}{(X-1)^4} + \frac{5}{(X-1)^3} + \frac{8}{(X-1)^2} + \frac{11}{X-1} + \frac{3}{(X-2)^2} - \frac{11}{X-2}.$$

Application

La transformée de Laplace intervient dans la résolution d'équations et de systèmes différentiels notamment en électricité, électronique et théorie du signal.

Nous rappelons que si f est une fonction nulle en dehors de $\mathbb{R}^+$, sa transformée de Laplace est la fonction qui à tout réel ou complexe p associe, sous condition d'existence de l'intégrale :

$$\mathcal{L}[f](p) = \int_0^{+\infty} f(t)e^{-pt} dt.$$

Nous proposons de montrer ici comment la décomposition en éléments simples peut intervenir dans la résolution d'une équation différentielle. Posons l'équation suivante :

$$y''(t) = 3y'(t) - 2y(t) + e^t$$

avec pour conditions initiales $y(0) = 1$ et $y'(0) = 0$. Cherchons la solution de cette équation différentielle.



Nous proposons tout d'abord, afin de résoudre cette équation différentielle, d'utiliser la transformée de Laplace en l'appliquant à notre équation ci-dessus pour obtenir, via les propriétés bien connues de la transformée de Laplace :

$$(p^2 - 3p + 2)\mathcal{L}[y](p) - p + 3 = \frac{1}{p-1}.$$

Nous en déduisons la décomposition en éléments simples suivante :

$$\begin{aligned}\mathcal{L}[y](p) &= \frac{p^2 - 4p + 4}{(p^2 - 3p + 2)(p-1)} \\ &= \frac{p-2}{(p-1)^2} \\ &= \frac{A}{p-1} + \frac{B}{(p-1)^2}.\end{aligned}$$



Il ne reste plus qu'à déterminer les coefficients A et B de la décomposition en éléments simples précédente.

Nous multiplions l'expression précédente par $(p-1)^2$ et posons $p = 1$ ce qui nous amène à $B = -1$. Pour trouver A , il suffit de multiplier l'expression par $p-1$ et de prendre la limite lorsque p tend vers $+\infty$. Nous trouvons ainsi $A = 1$.



Fiche 69

Enfin, nous terminons la résolution à partir de l'expression

$$\mathcal{L}[y](p) = \frac{1}{p-1} - \frac{1}{(p-1)^2}$$

par passage à l'inverse de la transformée de Laplace pour obtenir :

$$y(t) = e^t - te^t = (1-t)e^t.$$

Pour s'entraîner (solutions p. 510)

13.1 Donner la décomposition en éléments simples des fractions suivantes :

1. $F_1(X) = \frac{1}{X^2(X-1)}$

2. $F_2(X) = \frac{X}{(X+1)^2(X-1)^3}$

3. $F_3(X) = \frac{X}{(X+3)(X^2+X+1)}$

4. $F_4(X) = \frac{1}{X^2-1}$

5. $F_5(X) = \frac{X+3}{X^4-5X^2+4}$

6. $F_6(X) = \frac{X^4+X+1}{X^3-3X^2+2X}$

13.2 Nous considérons l'équation différentielle :

$$y''(t) - y'(t) = \sin(2t)$$

avec $y(0) = y'(0) = 1$.

1. Si nous posons $Y = \mathcal{L}[y]$ la transformée de Laplace de y , montrer que nous avons :

 Fiche 69

$$p^2Y - pY - p = \frac{2}{p^2+4}.$$

2. Montrer que dans ce cas, $Y = \frac{a}{p} + \frac{b}{p-1} + \frac{cp+d}{p^2+4}$ et déterminer les coefficients a, b, c et d .

3. Par inversion de la transformée de Laplace, montrer que nous arrivons à :

$$y(t) = \frac{7}{5}e^t + \frac{1}{10}\cos 2t - \frac{1}{5}\sin 2t - \frac{1}{2}.$$

Les systèmes d'équations linéaires sont très importants en algèbre linéaire. Leur présence croissante dans les applications est due aux progrès considérables de l'informatique ces dernières années. Ils sont par exemple utilisés dans le domaine des transports ou bien encore en chimie, dans le but d'équilibrer des équations.

1. Définitions

Définition 14.1

Une **équation linéaire** à p inconnues $x_1, \dots, x_p$ est une équation de la forme :

$$a_1 x_1 + a_2 x_2 + \dots + a_p x_p = b,$$

où b et les coefficients $a_1, \dots, a_p$ sont des nombres réels ou complexes.

Exemples

1. $2x_1 + 5x_2 = 7$ est une équation de type linéaire.
2. $2x_2 + 5x_2 = x_1 x_2$ n'est pas une équation de type linéaire à cause du terme $x_1 x_2$.

Définition 14.2

Un **système linéaire** (S) de n équations $L_1, L_2, \dots, L_n$ à p inconnues $x_1, x_2, \dots, x_p$ s'écrit sous la forme :

$$(S) \begin{cases} a_{11} x_1 + a_{12} x_2 + \dots + a_{1p} x_p = b_1 & L_1 \\ a_{21} x_1 + a_{22} x_2 + \dots + a_{2p} x_p = b_2 & L_2 \\ \vdots & \vdots \\ a_{n1} x_1 + a_{n2} x_2 + \dots + a_{np} x_p = b_n & L_n \end{cases}$$

Les $(a_{ij})_{1 \leq i \leq n, 1 \leq j \leq p}$ et $(b_j)_{1 \leq j \leq n}$ sont des nombres réels ou complexes. Résoudre le système (S) consiste à déterminer les p -uplets $(x_1; \dots; x_p)$ qui vérifient simultanément les équations $L_1; L_2; \dots; L_n$. Ces p -uplets sont appelés solutions du système (S) . Un système linéaire est dit **homogène** si les $(b_i)_{1 \leq i \leq n}$ sont nuls. Si au moins un des $(b_i)_{1 \leq i \leq n}$ est non nul, le système est dit **non homogène**.

Exemple

Le système $(S) \begin{cases} 2x_1 + x_2 - x_3 = 5 \\ x_1 - x_2 + x_3 = 3 \end{cases}$ est un système linéaire non homogène de deux équations et trois inconnues.

Définition 14.3

Deux systèmes linéaires ayant le même nombre d'inconnues sont dits **équivalents** s'ils ont le même ensemble de solutions.

2. Algorithme de résolution

L'existence de solutions d'un système d'équations linéaires dépend du nombre d'équations qui sont linéairement indépendantes, c'est-à-dire telles que l'une ne soit pas une combinaison linéaire des autres. Nous proposons ici de remplacer le système de départ par un système équivalent, plus facile à résoudre. Pour cela, nous introduisons la définition d'un système de type triangulaire.

Définition 14.4

Un système d'équations linéaires (S) à n équations et p inconnues est **triangulaire** si $a_{ij} = 0$ pour tout $(i; j)$ tel que $i > j$ ou si $a_{ij} = 0$ pour tout $(i; j)$ tel que $i < j$.

Exemple

Le système $(S) \begin{cases} 2x_1 + x_2 + x_3 = 5 \\ x_2 - x_3 = 2 \\ x_3 = 1 \end{cases}$ est un système linéaire de type triangulaire.

Pour obtenir un système triangulaire, la procédure systématique de résolution de système linéaire ou algorithme de résolution consiste à se servir du terme en x_1 dans la première équation pour éliminer les termes en x_1 des autres équations. Puis à se servir du terme en x_2 dans la deuxième équation pour éliminer les termes en x_2 des autres équations. Nous réitérons la même opération avec les autres inconnues jusqu'à obtenir un système équivalent qui sera très simple à résoudre.

Théorème 14.1

Les opérations élémentaires ci-dessous transforment un système linéaire en un système linéaire équivalent.

- Permuter deux équations (i.e. deux lignes) L_i et L_j , ce que nous codons par $L_i \longleftrightarrow L_j$.
- Multiplier une équation L_i par un nombre λ réel ou complexe non nul, ce que nous codons par $L_i \longleftarrow \lambda L_i$ avec $\lambda \neq 0$.
- Ajouter à la ligne L_i un multiple de la ligne L_j avec $j \neq i$, ce que nous codons par $L_i \longleftarrow L_i + \lambda L_j$, avec $\lambda \in \mathbb{R}$ ou $\mathbb{C}$.

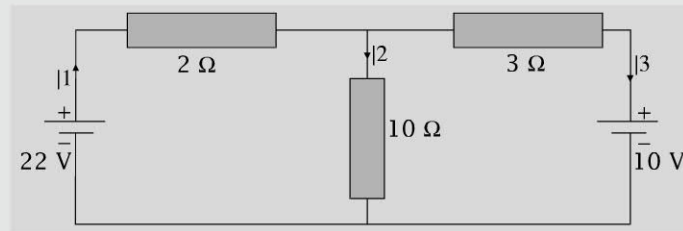


Les trois types d'ensembles de solutions possibles sont l'ensemble vide, un singleton, associé à une solution unique, ou un ensemble contenant un nombre infini de solutions.

Dans les Fiches 15 et 25, nous présenterons d'autres méthodes de résolution des systèmes linéaires telles que la méthode du pivot de Gauss ou encore celle qui fait appel au calcul de déterminants.

Application

Nous considérons le circuit électrique suivant.



Calculer les intensités des trois courants I_1 , I_2 et I_3 qui circulent dans les trois branches du circuit ci-dessus sachant qu'ils sont définis par les équations :

$$\begin{aligned} I_1 &= I_2 + I_3, \\ 2I_1 &= 22 - 10I_2, \\ 10I_2 &= 3I_3 + 10. \end{aligned}$$

Ce système est un système linéaire non homogène de trois équations à trois inconnues de la forme :

$$(S) \begin{cases} I_1 - I_2 - I_3 = 0 & L_1 \\ 2I_1 + 10I_2 = 22 & L_2 \\ 10I_2 - 3I_3 = 10 & L_3 \end{cases}$$



Nous utilisons les opérations successives $L_2 \leftarrow L_2 - 2L_1$ puis $L_2 \leftarrow \frac{1}{2}L_2$ et $L_3 \leftarrow L_3 - \frac{5}{3}L_2$ pour le transformer en un système triangulaire équivalent.

$$(S') \begin{cases} I_1 - I_2 - I_3 = 0 & L_1 \\ 12I_2 + 2I_3 = 22 & L_2 \leftarrow L_2 - 2L_1 \\ 10I_2 - 3I_3 = 10 & L_3 \end{cases}$$

puis

$$(S'') \begin{cases} I_1 - I_2 - I_3 = 0 & L_1 \\ 6I_2 + I_3 = 11 & L_2 \leftarrow \frac{1}{2}L_2 \\ 10I_2 - 3I_3 = 10 & L_3 \end{cases}$$

et enfin

$$(S''') \begin{cases} I_1 - I_2 - I_3 = 0 & L_1 \\ 6I_2 + I_3 = 11 & L_2 \\ -\frac{14}{3}I_3 = -\frac{25}{3} & L_3 \leftarrow L_3 - \frac{5}{3}L_2 \end{cases}$$

Nous obtenons ainsi un système (S''') triangulaire et équivalent au système (S) de départ. Nous commençons par résoudre la dernière équation pour trouver I_3 . Nous obtenons $I_3 = \frac{25}{14}$. Puis nous injectons le résultat obtenu pour I_3 dans la deuxième équation pour obtenir $I_2 = \frac{11-I_3}{6} \approx 1,5A$. Enfin, en injectant les résultats de I_2 et I_3 dans la première équation nous trouvons $I_1 = I_2 + I_3 \approx 3,3A$.

La solution du système (S) est donc le triplet $\{(3,3; 1,5; 1,8)\}$.

Pour s'entraîner (solutions p. 510)

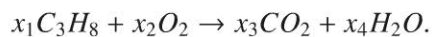
14.1 Résoudre le système homogène suivant :

$$(S) \quad 5x - 2y + 3z = 0.$$

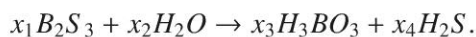
14.2 Résoudre le système homogène suivant :

$$(S) \quad \begin{cases} 3x + 5y - 4z = 0 \\ -3x - 2y + 4z = 0 \\ 6x + y - 8z = 0 \end{cases}$$

14.3 Pondérer l'équation chimique suivante :



14.4 Pondérer l'équation chimique suivante :



14.5 Résoudre les trois systèmes non homogènes suivants :

$$1. (S_1) \quad \begin{cases} -x + y - 2z = -4 \\ -x - y - z = -11 \\ 2x + y - z = 8 \end{cases}$$

$$2. (S_2) \quad \begin{cases} 2x + y - z = 1 \\ x - y + z = 2 \\ 4x + 3y + z = 3 \end{cases}$$

$$3. (S_3) \quad \begin{cases} x + 2y + z + t = 0 \\ x + y - z - t = 8 \\ 2x + y + z + t = -1 \\ -x + y + z - 2t = -2 \end{cases}$$

La méthode du pivot de Gauss permet de transformer facilement un système linéaire en un autre système équivalent mais plus simple à résoudre. Elle est de ce fait très utilisée en algèbre linéaire et dans les applications.

1. Matrice échelonnée, augmentée

Définition 15.1

Un tableau rectangulaire, appelé matrice, suffit à résumer l'essentiel de l'information présente dans un système linéaire. Étant donné le système linéaire (S) :

$$(S) \begin{cases} a_{11} x_1 + a_{12} x_2 + \cdots + a_{1p} x_p = b_1 \\ a_{21} x_1 + a_{22} x_2 + \cdots + a_{2p} x_p = b_2 \\ \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \\ a_{n1} x_1 + a_{n2} x_2 + \cdots + a_{np} x_p = b_n \end{cases},$$

la **matrice des coefficients** de (S) est définie de telle façon que les coefficients de chaque inconnue soient disposés en colonnes :

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{pmatrix}.$$

Nous appelons **matrice augmentée**, la matrice des coefficients de (S) à laquelle nous ajoutons la colonne des termes des membres de droite de (S) :

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1p} & b_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} & b_n \end{array} \right).$$

Définition 15.2

Une matrice rectangulaire est dite sous forme **échelonnée** si elle remplit les trois conditions suivantes :

- toutes ses lignes non nulles sont situées au-dessus de ses lignes nulles,
- chaque élément de tête de ligne se trouve dans une colonne à droite de l'élément de tête de ligne précédente,

- tous les éléments de la colonne sous un élément de tête sont nuls.

Elle est dite sous forme **échelonnée réduite** si en plus d'être échelonnée elle satisfait aux conditions supplémentaires suivantes :

- L'élément de tête de chaque ligne non nulle vaut 1,
- Chaque 1 de tête d'une ligne est le seul élément non nul de sa colonne.

Exemple

Les matrices $\begin{pmatrix} 2 & 1 & 3 & 3 \\ 0 & 3 & 0 & 6 \\ 0 & 0 & 5 & 3 \end{pmatrix}$ et $\begin{pmatrix} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & 6 \\ 0 & 0 & 1 & 3 \end{pmatrix}$ sont respectivement sous forme échelonnée et échelonnée réduite.

2. Pivot de Gauss

Définition 15.3

Une **position pivot** dans une matrice est celle qu'occupe un 1 de tête dans la forme échelonnée réduite de cette matrice. Une colonne qui contient une position pivot est appelée **colonne pivot**.

Théorème 15.1

Les opérations qu'il est possible d'effectuer sur les lignes dans le cadre d'une résolution d'un système par la méthode du pivot sont les suivantes :

- remplacer une ligne par la somme de cette ligne avec un multiple d'une autre ligne,
- échanger deux lignes,
- multiplier tous les éléments d'une ligne par une constante non nulle.

► Algorithme du Pivot de Gauss

Nous travaillons sur la matrice augmentée obtenue à partir du système d'équations linéaires de départ.

- Nous commençons par la colonne non nulle la plus à gauche. C'est une colonne pivot. La position pivot est en haut de la colonne.
- Nous choisissons comme pivot un élément non nul de la colonne pivot. Il est parfois nécessaire d'invertir les lignes pour que cet élément occupe la position pivot.
- Nous appliquons les opérations élémentaires sur les lignes pour que tous les éléments sous le pivot deviennent des 0.
- Nous abandonnons la ligne qui contient le pivot ainsi que toutes les lignes au dessus de celle-ci s'il y en a. Nous appliquons à nouveau les étapes précédentes à la sous-matrice qu'il reste. Nous répétons ceci jusqu'à ce qu'il n'y ait plus de ligne non nulle à changer.
- Finalement, nous réécrivons sous forme d'équations le système correspondant à la matrice de l'étape précédente. Ce système est équivalent au système de départ.

- Lorsque l'ensemble des solutions $\mathcal{S}$ est infini, il est nécessaire de considérer des variables libres et liées afin de pouvoir paramétrer $\mathcal{S}$ et décrire toutes les solutions.



Si nous débutons avec un pivot le plus à droite et en travaillant de bas en haut vers la gauche, nous faisons apparaître des zéros au dessus de chaque pivot. Si un pivot n'est pas égal à 1, nous le faisons devenir égal à 1 par une des opérations précédemment citées.

Application

Résoudre le système d'équations linéaires suivant en utilisant la méthode du Pivot de Gauss :

$$(S) \begin{cases} x + 2y + 2z = 2 \\ x + 3y - 2z = -1 \\ 3x + 5y + 8z = 8 \end{cases}$$

La matrice augmentée associée à (S) est :

$$\left(\begin{array}{ccc|c} 1 & 2 & 2 & 2 \\ 1 & 3 & -2 & -1 \\ 3 & 5 & 8 & 8 \end{array} \right).$$

Nous choisissons 1 en haut à gauche du tableau comme premier pivot et à partir des opérations

$$L_2 \leftarrow L_2 - L_1 \text{ et } L_3 \leftarrow L_3 - 3L_1$$

sur les lignes du tableau (où L_1, L_2 et L_3 désignent respectivement les première, deuxième et troisième lignes du tableau), nous faisons apparaître des 0 en dessous pour obtenir :

$$\left(\begin{array}{ccc|c} 1 & 2 & 2 & 2 \\ 0 & 1 & -4 & -3 \\ 0 & -1 & 2 & 2 \end{array} \right).$$

Nous choisissons ensuite un deuxième pivot sur la deuxième ligne et nous effectuons les opérations

$$L_1 \leftarrow L_1 - 2L_2 \text{ et } L_3 \leftarrow L_3 + L_2$$

pour obtenir :

$$\left(\begin{array}{ccc|c} 1 & 0 & 10 & 8 \\ 0 & 1 & -4 & -3 \\ 0 & 0 & -2 & -1 \end{array} \right).$$

En divisant la dernière ligne par -2 puis en effectuant les opérations

$$L_1 \leftarrow L_1 - 10L_3 \text{ et } L_2 \leftarrow L_2 + 4L_3$$

nous obtenons la matrice échelonnée réduite suivante :

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & \frac{1}{2} \end{array} \right).$$

Cela nous permet de déduire directement que $x = 3$, $y = -1$ et $z = \frac{1}{2}$.

La solution du système (S) est alors $\{(3; -1; \frac{1}{2})\}$.

Pour s'entraîner
(solutions p. 511)

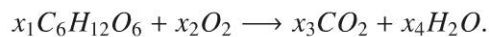
15.1 Résoudre le système suivant avec la méthode du Pivot de Gauss :

$$(S) \begin{cases} x + y - z = 1 \\ 2x - 3y + 2z = -2 \\ 3x + y - z = 3 \end{cases}$$

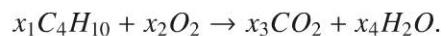
15.2 Résoudre le système suivant en utilisant la méthode du Pivot de Gauss :

$$(S) \begin{cases} -x + y - 2z = -4 \\ -x - y - z = -11 \\ 2x + y - z = 8 \end{cases}$$

15.3 Pondérer l'équation chimique suivante qui correspond à la dégradation de la molécule du glucose ($C_6H_{12}O_6$) dans l'organisme :



15.4 Pondérer l'équation chimique suivante qui correspond à la combustion du butane :



15.5 Résoudre les deux systèmes non homogènes suivants à partir de la méthode du Pivot de Gauss :

$$1. (S_1) \begin{cases} 2x + y - z = 1 \\ x - y + z = 2 \\ 4x + 3y + z = 3 \end{cases}$$

$$2. (S_2) \begin{cases} x - y + 2z = 5 \\ 3x + 2y + z = 10 \\ 2x - 3y - 2z = -10 \end{cases}$$

Nombres complexes

L'ensemble $\mathbb{C}$ des nombres complexes est une extension de l'ensemble des nombres réels. En analyse, ils sont utilisés pour l'étude des séries de Fourier. En Physique, ils sont utilisés pour décrire le comportement d'oscillateurs électriques ou les phénomènes ondulatoire en électromagnétisme. L'équation $x^2 + 1 = 0$ n'a pas de solution réelle, mais possède deux solutions dans l'ensemble des nombres complexes. Ces deux solutions sont appelées **racines carrées de -1** et sont notées i et $-i$. Nous avons $i^2 = (-i)^2 = -1$.

1. Forme algébrique

Définition 16.1

Soit $z \in \mathbb{C}$. Alors, l'écriture de z sous **forme algébrique** est donnée par l'unique couple $(x; y)$ de nombres réels tel que $z = x + iy$. Le réel x est appelé **partie réelle** de z et noté $\Re(z)$. Le réel y est appelé **partie imaginaire** de z et noté $\Im(z)$. Nous appelons nombre complexe **conjugué** de z le nombre complexe $\bar{z} = x - iy$. Un nombre **imaginaire pur** est un nombre complexe qui s'écrit sous la forme ai avec $a \in \mathbb{R}$.



L'écriture algébrique d'un nombre complexe est unique. En particulier, soient x, y, x', y' des nombres réels, et $z = x + iy$ et $z' = x' + iy'$ deux nombres complexes. Nous avons $z = z' \iff (x = x' \text{ et } y = y')$.

Exemple

$z = 2 + 3i$ est un nombre complexe avec la partie réelle égale à 2 et la partie imaginaire égale à 3. Le complexe conjugué de z est $\bar{z} = 2 - 3i$.

Théorème 16.1

Soient z et z' (z' non nul) des nombres complexes. Nous avons les propriétés suivantes :

$$\bar{\bar{z}} = z, \quad \overline{z + z'} = \bar{z} + \bar{z'}, \quad \overline{zz'} = \bar{z}\bar{z'}, \quad \overline{\left(\frac{z}{z'}\right)} = \frac{\bar{z}}{\bar{z'}}.$$

Théorème 16.2

L'ensemble $(\mathbb{C}, +, \times)$ est un corps. En particulier, soient x, y, x', y' des nombres réels, et $z = x + iy$ et $z' = x' + iy'$ deux nombres complexes. Nous avons

$$z + z' = (x + x') + i(y + y') \text{ et } zz' = (xx' - yy') + i(xy' + x'y).$$

2. Formes trigonométrique et exponentielle

Définition 16.2

- Soit $z = x + iy$ un nombre complexe avec $x, y \in \mathbb{R}$. Alors, le **module** de z est le réel positif, noté $|z|$, défini par

$$|z| = \sqrt{x^2 + y^2} = \sqrt{z\bar{z}}.$$

- Tout nombre complexe z non nul s'écrit sous la **forme trigonométrique**

$$z = \rho(\cos \theta + i \sin \theta) \text{ avec } \rho > 0 \text{ et } \theta \in \mathbb{R}$$

où $\rho = |z|$ est le module de z et θ est un argument de z , noté $\arg(z)$ et défini modulo 2π par $\cos \theta = \frac{x}{\rho}$ et $\sin \theta = \frac{y}{\rho}$.



En général, expliciter l'argument d'un nombre complexe n'est pas possible.

Définition 16.3

Pour tout $\theta \in \mathbb{R}$, nous appelons **exponentielle de l'imaginaire pur** $i\theta$ le nombre complexe défini par :

$$e^{i\theta} = \cos(\theta) + i \sin(\theta).$$

Tout nombre complexe z non nul s'écrit alors sous sa **forme exponentielle** : $z = \rho e^{i\theta}$ où $\rho = |z|$ est le module de z et $\theta = \arg(z)$ est un argument de z .

Théorème 16.3

- **Formule de Moivre.** $\forall \theta \in \mathbb{R}, \forall n \in \mathbb{Z}, (\cos \theta + i \sin \theta)^n = \cos n\theta + i \sin n\theta$, ce qui s'écrit, avec la notation exponentielle : $(e^{i\theta})^n = e^{in\theta}$.
- **Formules d'Euler.** $\forall \theta \in \mathbb{R}, \cos \theta = \frac{e^{i\theta} + e^{-i\theta}}{2}$ et $\sin \theta = \frac{e^{i\theta} - e^{-i\theta}}{2i}$.



Les formules d'Euler permettent de linéariser les polynômes trigonométriques. Ceci est utile pour calculer des primitives ou des intégrales dans lesquelles interviennent ces polynômes.

3. Racines n -ièmes d'un nombre complexe

Définition 16.4

- Soit $n \in \mathbb{N}^*$. Le nombre complexe z est une **racine** n -ième de l'unité si $z^n = 1$.
- Soient $n \in \mathbb{N}^*$ et $z_0 \in \mathbb{C}^*$. Le nombre complexe z est une **racine** n -ième de z_0 si $z^n = z_0$.

Théorème 16.4

- Soit $n \in \mathbb{N}^*$. Les racines n -ièmes de l'unité sont les n complexes ω_k , définis pour $k \in \{0; \dots; n-1\}$, par :

$$\omega_k = e^{i\frac{2k\pi}{n}} = \cos \frac{2k\pi}{n} + i \sin \frac{2k\pi}{n}.$$

- Soit $n \in \mathbb{N}^*$. Les racines n -ièmes d'un nombre complexe non nul $z_0 = \rho_0(\cos \theta_0 + i \sin \theta_0) = \rho_0 e^{i\theta_0}$ sont les n complexes, pour $k \in \{0; \dots; n-1\}$,

$$z_k = \sqrt[n]{\rho_0} e^{i\frac{\theta_0}{n} + i\frac{2ik\pi}{n}}.$$



En général, nous ne savons pas extraire de racines n -ièmes d'un nombre puisque nous ne pouvons pas déterminer explicitement son argument. Ceci nous amène à considérer une autre méthode pour calculer les racines carrées.

Théorème 16.5 Cas particulier des racines carrées

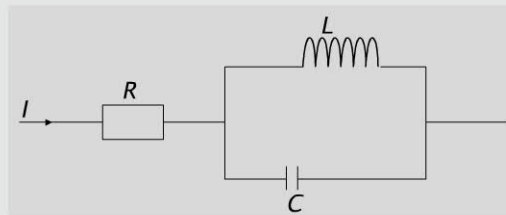
Soient a, b deux réels. Pour trouver le nombre complexe $z = x + iy$ solution de $z^2 = a + ib$, nous procédons par identification. En effet, nous avons :

- $x^2 - y^2 = a$ (égalité des parties réelles),
- $2xy = b$ (égalité des parties imaginaires),
- $x^2 + y^2 = \sqrt{a^2 + b^2}$ (égalité des modules),

ce qui permet de déduire x et y . Noter que ce calcul est utile lorsque nous souhaitons résoudre une équation du second degré à coefficients complexes.

Application

L'impédance électrique mesure l'opposition d'un circuit électrique au passage d'un courant alternatif sinusoïdal. Nous considérons le circuit ci-dessous alimenté par un courant sinusoïdal I où R désigne une résistance, C un condensateur et L une bobine.



1. Calculer l'impédance complexe de ce circuit.
2. Pour quelle fréquence le courant I est-il nul ?

1.



Donnons d'abord les impédances complexes de chacun des dipôles passifs élémentaires qui composent ce circuit. Si ω définit la pulsation (en radians par seconde), l'impédance complexe

- d'un condensateur C est donnée par $Z_C = \frac{1}{iC\omega}$, où C est la capacité (en farad F) du condensateur ;
- d'une bobine L est donnée par $Z_L = iL\omega$, où L est l'inductance (en henry H) de la bobine ;
- d'une résistance R est donnée par $Z_R = R$, où R est la valeur (en ohms Ω) de la résistance.

Pour déterminer l'impédance complexe du circuit, il faut savoir qu'en série les impédances complexes s'additionnent alors qu'en parallèle ce sont les admittances complexes qui s'additionnent.

Compte tenu de ces remarques, l'impédance du circuit est :

$$Z_{tot} = Z_R + \frac{1}{\frac{1}{Z_L} + \frac{1}{Z_C}} = R + \frac{1}{\frac{1}{iL\omega} + \frac{1}{\frac{1}{iC\omega}}} = R + \frac{iL\omega \times \frac{1}{iC\omega}}{iL\omega + \frac{1}{iC\omega}} = R - i \frac{L\omega}{LC\omega^2 - 1}.$$

2. Le courant I est nul lorsque l'impédance est infinie. Si nous reprenons l'expression de l'impédance complexe obtenue à la question 1, I est nul lorsque $LC\omega^2 - 1 = 0$ ou encore $LC\omega^2 = 1$. Comme de plus $\omega = 2\pi f$ où f désigne la fréquence, nous en déduisons que le courant est nul pour la fréquence :

$$f = \frac{1}{2\pi\sqrt{LC}}.$$

Pour s'entraîner (solutions p. 511)

16.1 Nous considérons les nombres complexes $z = 3 + 5i$ et $z' = -1 + 2i$. Écrire sous forme algébrique $z_1 = z - \overline{z'}$, $z_2 = z^2$ et $z_3 = z\overline{z}$.

16.2 Nous considérons les deux nombres complexes $z_1 = e^{i\frac{\pi}{3}}$ et $z_2 = e^{-i\frac{\pi}{4}}$.

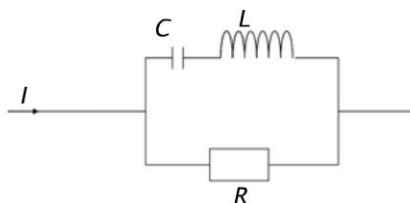
1. Écrire z_1 et z_2 sous forme algébrique.
2. Donner les écritures sous forme algébrique, exponentielle et trigonométrique de $z_1 z_2$.
3. En déduire la valeur exacte de $\cos \frac{\pi}{12}$ et $\sin \frac{\pi}{12}$.

16.3 Résoudre dans $\mathbb{C}$ les équations suivantes :

1. $5z + 2i = (1 + i)z - 3$.
2. $\frac{z - i}{z + 1} = 4i$.

16.4 Soit z un nombre complexe tel que $1 + z^4 + z^8 = 0$. Montrer que z est racine douzième de l'unité.

16.5 Nous proposons l'étude du circuit ci-dessous en régime sinusoïdal :



où R désigne une résistance, C un condensateur et L une bobine. I représente le courant parcouru par ce circuit.

1. Calculer l'impédance complexe de ce circuit.
2. Pour quelle fréquence l'impédance est-elle nulle ?

16.6 Utiliser les formules d'Euler pour montrer que, pour tout $\theta \in \mathbb{R}$:

$$(\sin \theta)^2 = \frac{1 - \cos(2\theta)}{2}.$$

Nombres complexes et géométrie plane

1. Définitions

Définition 17.1

Soit le repère orthonormal $(O; \vec{i}; \vec{j})$ du plan. Pour tout couple $(x; y)$ de réels, $M \begin{pmatrix} x \\ y \end{pmatrix}$ désigne le point du plan de coordonnées $(x; y)$. L'application de $\mathbb{C}$ dans $\mathbb{R}^2$ qui à tout nombre complexe z associe le point M de coordonnées $(x; y)$ est une bijection. M est appelé **image** de z , z est appelé **affixe** du point M . Le vecteur $\vec{u} = \overrightarrow{OM}$ est l'**image vectorielle** de z et a pour affixe z .



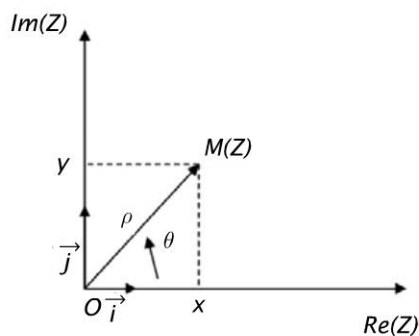
Les images des nombres complexes z et $\bar{z}$ sont symétriques par rapport à l'axe des abscisses.

Définition 17.2

- Si M est l'image de z dans le repère orthonormal $(O; \vec{i}; \vec{j})$ du plan, alors le module $|z|$ de z , représente la longueur OM .
- Tout nombre complexe non nul d'image M dans le repère orthonormal $(O; \vec{i}; \vec{j})$ du plan s'écrit sous la forme trigonométrique suivante :

$$z = \rho(\cos \theta + i \sin \theta),$$

avec ρ et θ respectivement le module et l'argument de z . L'argument θ est l'unique mesure en radians de l'angle orienté $(\vec{u}; \overrightarrow{OM})$ dans $[0; 2\pi[$.



2. Transformations géométriques

Dans la suite, nous supposons que $z_1 = x_1 + iy_1$ et $z_2 = x_2 + iy_2$ sont deux nombres complexes donnés.

Définition 17.3

- L'application de $\mathbb{C}$ dans $\mathbb{C}$ qui à z associe $z + z_2$ se traduit sur les images par la **translation** de vecteur $x_2 \vec{i} + y_2 \vec{j}$.
- Soit $\theta \in [0; 2\pi[$. L'application de $\mathbb{C}$ dans $\mathbb{C}$ qui à z associe $e^{i\theta}z$ se traduit sur les images par la **rotation** de centre O l'origine du repère et d'angle θ .
- Si $z_1 \neq 1$, l'application de $\mathbb{C}$ dans $\mathbb{C}$ qui à z associe $z_1 z + z_2$ se traduit sur les images par la **similitude** de rapport $|z_1|$, d'angle $\arg(z_1)$ et dont le centre Ω a pour affixe $z_\Omega = \frac{1 - z_2}{z_1}$.
Cette transformation est la composée, dans n'importe quel ordre, de la rotation de centre Ω et d'angle $\arg(z_1)$ et de l'homothétie de centre Ω et de rapport $|z_1|$.

3. Applications géométriques

Théorème 17.1

Soient M_1 , M_2 et M_3 , trois points distincts d'affixes respectifs z_{M_1} , z_{M_2} et z_{M_3} . Alors, nous avons :

- le module $|z_{M_2} - z_{M_1}|$ est la longueur $M_1 M_2$ et l'argument $\arg(z_{M_2} - z_{M_1})$ est une mesure de l'angle $(\vec{i}; \overrightarrow{M_1 M_2})$.
- $\frac{z_{M_2} - z_{M_1}}{z_{M_3} - z_{M_1}}$ a pour module $\frac{M_1 M_2}{M_1 M_3}$ et pour argument $\arg(z_{M_2} - z_{M_1}) - \arg(z_{M_3} - z_{M_1})$, qui est une mesure de l'angle $(\overrightarrow{M_1 M_3}, \overrightarrow{M_1 M_2})$.

Théorème 17.2

Soient M_1 , M_2 et M_3 , trois points distincts d'affixes respectifs z_{M_1} , z_{M_2} et z_{M_3} . Alors, nous avons :

- Les points M_1 , M_2 et M_3 sont alignés si et seulement si $\frac{z_{M_2} - z_{M_1}}{z_{M_3} - z_{M_1}}$ est un réel.
- $\overrightarrow{M_1 M_2}$ et $\overrightarrow{M_1 M_3}$ sont orthogonaux si et seulement si $\frac{z_{M_2} - z_{M_1}}{z_{M_3} - z_{M_1}}$ est un imaginaire pur.

Théorème 17.3

Soient M_1 , M_2 et M_3 , trois points distincts d'affixes respectifs z_{M_1} , z_{M_2} et z_{M_3} . Alors, nous avons :

- Le triangle $M_1 M_2 M_3$ est équilatéral, de sens direct, si et seulement si

$$z_{M_3} - z_{M_1} = e^{i\frac{\pi}{3}}(z_{M_2} - z_{M_1}).$$

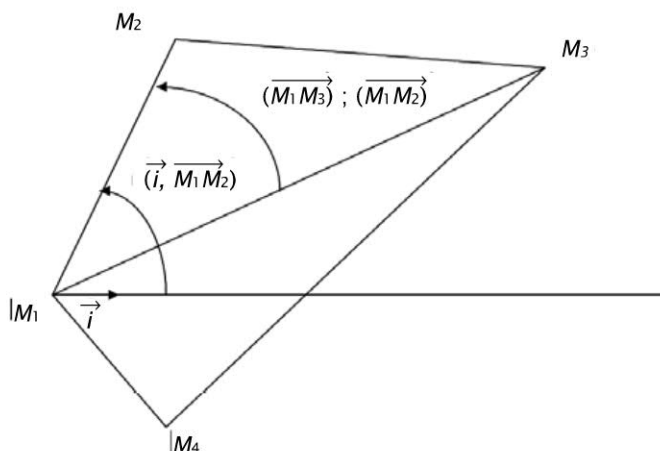
- Le triangle $M_1 M_2 M_3$ est isocèle et rectangle en M_1 si et seulement si

$$z_{M_3} - z_{M_1} = -i(z_{M_2} - z_{M_1}) \text{ ou } z_{M_3} - z_{M_1} = i(z_{M_2} - z_{M_1}).$$

Théorème 17.4

Soient M_1, M_2, M_3 et M_4 quatre points, deux à deux distincts, d'affixes respectifs $z_{M_1}, z_{M_2}, z_{M_3}$ et z_{M_4} . Ils sont cocycliques ou alignés si et seulement si

$$\frac{\left(\frac{z_{M_2}-z_{M_3}}{z_{M_1}-z_{M_3}}\right)}{\left(\frac{z_{M_2}-z_{M_4}}{z_{M_1}-z_{M_4}}\right)} \in \mathbb{R}.$$



Application

Nous considérons les points A, B et C d'affixes respectifs $z_A = 3, z_B = \frac{5}{2} + \frac{7}{2}i$ et $z_C = -\frac{1}{2} - \frac{1}{2}i$.

1. Montrer que le triangle ABC est rectangle et isocèle.
2. Donner l'affixe du point I milieu de $[BC]$.
3. Donner l'affixe du point D afin que $ABDC$ soit un carré.

1. Montrons que le triangle ABC est rectangle et isocèle en A .

- $AB = |z_B - z_A| = \left|\frac{5}{2} + \frac{7}{2}i - 3\right| = \left|-\frac{1}{2} + \frac{7}{2}i\right| = \sqrt{\left(-\frac{1}{2}\right)^2 + \left(\frac{7}{2}\right)^2} = \sqrt{\frac{50}{4}} = \frac{5}{2}\sqrt{2}.$
- $AC = |z_C - z_A| = \left|-\frac{1}{2} - \frac{1}{2}i - 3\right| = \left|-\frac{7}{2} - \frac{1}{2}i\right| = \sqrt{\left(-\frac{7}{2}\right)^2 + \left(-\frac{1}{2}\right)^2} = \sqrt{\frac{50}{4}} = \frac{5}{2}\sqrt{2}.$
- $BC = |z_C - z_B| = \left|-\frac{1}{2} - \frac{1}{2}i - \frac{5}{2} - \frac{7}{2}i\right| = \left|-3 - 4i\right| = \sqrt{9 + 16} = 5.$

Nous en déduisons donc que ABC est isocèle en A puisque $AB = AC$.

De plus, $BC^2 = 25 = \frac{25}{2} + \frac{25}{2} = AB^2 + AC^2$, donc d'après le théorème de Pythagore, ABC est rectangle en A .



Autre raisonnement : $z_C - z_A = -\frac{7}{2} - \frac{1}{2}i$ et $z_B - z_A = -\frac{1}{2} + \frac{7}{2}i$.

Nous en déduisons que $z_C - z_A = i(z_B - z_A)$ et donc, d'après la partie cours, le triangle ABC est rectangle et isocèle en A .

2. I est le milieu de $[BC]$ donc $z_I = \frac{z_B + z_C}{2} = 1 + \frac{3}{2}i$.

3. Cherchons $z_D = x_D + iy_D$, l'affixe du point D tel que $ABDC$ soit un carré. Si $ABDC$ est un carré alors $\overrightarrow{AB} = \overrightarrow{CD}$ ou encore $z_B - z_A = z_D - z_C$.

Nous devons donc résoudre l'équation :

$$-\frac{1}{2} + \frac{7}{2}i = x_D + \frac{1}{2} + \left(y_D + \frac{1}{2}\right)i,$$

ce qui nous amène à $-\frac{1}{2} = x_D + \frac{1}{2}$ et $\frac{7}{2} = y_D + \frac{1}{2}$ qui ont pour solution $x_D = -1$ et $y_D = 3$.

Le point D a donc pour affixe $z_D = -1 + 3i$.

Pour s'entraîner (solutions p. 511)

17.1 Soient les points A et B d'affixes respectifs : $z_A = 3 + 3i$ et $z_B = 3 - 3i$.

1. Calculer les modules de z_A et z_B . Que pouvons-nous en déduire sur la nature du triangle OAB ?

2. Calculer les arguments de z_A et z_B . Que pouvons-nous en déduire sur l'angle $(\overrightarrow{OA}; \overrightarrow{OB})$?

3. Que pouvons-nous conclure des questions précédentes sur le triangle OAB ?

17.2 Soient A, B et C d'affixes respectives a, b et c , et $j = e^{i\frac{2\pi}{3}}$. Montrer que le triangle ABC est équilatéral direct si et seulement si $a + cj + bj^2 = 0$.

17.3 Déterminer l'ensemble des points M d'affixe z tels que :

1. $|z - 3i| = 5$.

2. $z \cdot \bar{z} = 4$.

3. $\left| \frac{z - z_A}{z - z_B} \right| = 1$ où les points A et B d'affixes respectives z_A et z_B sont supposés connus.

17.4 Montrer que les points A, B et C sont alignés si et seulement si $\frac{z_B - z_A}{z_C - z_A}$ est un réel.

17.5 Montrer que les vecteurs $\overrightarrow{AB}$ et $\overrightarrow{AC}$ sont orthogonaux si et seulement si $\frac{z_B - z_A}{z_C - z_A}$ est un imaginaire pur.

Dans toute cette fiche, $\mathbb{K}$ désigne un corps.

1. Espaces vectoriels, sous-espaces vectoriels

Définition 18.1

Soit E un ensemble. Supposons qu'il existe sur E :

- une addition (loi interne) c'est-à-dire $\forall u, v \in E, u + v \in E$,
- une multiplication (loi externe) c'est-à-dire $\forall u \in E, \forall \lambda \in \mathbb{K}, \lambda u \in E$.

Alors E est un $\mathbb{K}$ -**espace vectoriel** si :

- l'addition est associative : $\forall u, v, w \in E, u + (v + w) = (u + v) + w$.
- l'addition est commutative : $\forall u, v \in E, u + v = v + u$.
- l'addition possède un élément neutre (0_E) : $\forall u \in E, u + 0_E = u$.
- $\forall u \in E$, il existe son opposé $-u$ tel que $u + (-u) = 0_E$.
- La multiplication est associative : $\forall \lambda, \mu \in \mathbb{K}, \forall u \in E, \lambda(\mu u) = (\lambda\mu)u$.
- la multiplication est distributive par rapport à l'addition :
 $\forall \lambda \in \mathbb{K}, \forall u, v \in E, \lambda(u + v) = \lambda u + \lambda v$.
- $\forall \lambda, \mu \in \mathbb{K}, \forall u \in E, (\lambda + \mu)u = \lambda u + \mu u$.
- $\forall u \in E, 1_{\mathbb{K}}.u = u$. $1_{\mathbb{K}}$ est appelé l'élément neutre.

Exemples

1. Soit $n \in \mathbb{N}^*$, $\mathbb{K}^n$ est un $\mathbb{K}$ -espace vectoriel. Par exemple, $\mathbb{R}^2$ et $\mathbb{R}^3$ et plus généralement $\mathbb{R}^n, n \in \mathbb{N}^*$ sont des $\mathbb{R}$ -espaces vectoriels,
2. $\mathcal{C}^n(\mathbb{R}; \mathbb{R})$, l'ensemble des fonctions de $\mathbb{R}$ dans $\mathbb{R}$ de classe $\mathcal{C}^n$ sur $\mathbb{R}$ est un $\mathbb{R}$ -espace vectoriel,
3. L'ensemble $\mathbb{R}_n[X]$ des polynômes de degré inférieur ou égal à n est un $\mathbb{R}$ -espace vectoriel.

 Fiche 37

Définition 18.2

Soient E un $\mathbb{K}$ -espace vectoriel et F un sous-ensemble de E . Alors F est un **sous-espace vectoriel** (sev) de E si et seulement si :

- $F \neq \emptyset$ (F est non vide).
- $\forall u, v \in F, u + v \in F$.
- $\forall \lambda \in \mathbb{K}, \forall u \in F, \lambda u \in F$.



Point méthode

- Pour montrer que F est un sous-espace vectoriel d'un $\mathbb{K}$ -espace vectoriel E : il suffit de montrer que F est non vide et que $\forall \lambda, \beta \in \mathbb{K}, \forall u, v \in F, \lambda u + \beta v \in F$.
- Pour montrer que F est non vide, il suffit de vérifier que $0_E \in F$.

Exemple

$F = \{(x; y) \in \mathbb{R}^2 ; x + 3y = 0\}$ est un sous-espace vectoriel de $\mathbb{R}^2$. En effet,

- $(0; 0) \in F$ puisque $0 + 3 \times 0 = 0$. Donc F est non vide.
- Soient $u = (x; y) \in F, v = (x'; y') \in F$ et $\lambda \in \mathbb{R}$, alors

$$(x + x') + 3(y + y') = x + 3y + x' + 3y' = 0,$$

$$\lambda x + 3(\lambda y) = \lambda(x + 3y) = 0.$$

Nous en déduisons que $u + v \in F$ et $\lambda u \in F$.

F est donc bien un sous-espace vectoriel de $\mathbb{R}^2$.

2. Somme, somme directe et espaces supplémentaires

Définition 18.3

Soient E un $\mathbb{K}$ -espace vectoriel et F et G , deux sous-espaces vectoriels de E . Alors, l'ensemble

$$F + G = \{x \in E ; \exists (x_1; x_2) \in F \times G, x = x_1 + x_2\},$$

définit la **somme** de F et G . C'est de plus un sous-espace vectoriel de E .

Définition 18.4

Soient E un $\mathbb{K}$ -espace vectoriel, F et G deux sous-espaces vectoriels de E . F et G sont en **somme directe**, notée $F \oplus G$, si et seulement si l'intersection de F et de G est égale à l'élément neutre : $F \cap G = \{0_E\}$.

Exemple

Soient $E = \mathbb{R}^3$ et deux sous-espaces vectoriels $F = \{(x; y; z) \in \mathbb{R}^3 ; z = 0\}$ et $G = \{(x; y; z) \in \mathbb{R}^3 ; x = y = 0\}$. Alors $F \cap G = \{0_{\mathbb{R}^3}\}$ donc F et G sont en somme directe.

Théorème 18.1

Pour que deux sous-espaces vectoriels F et G d'un $\mathbb{K}$ -espace vectoriel E soient en somme directe il suffit que tout élément x_E de $F + G$ se décompose de façon unique en somme d'un élément x_F de F et d'un élément x_G de G : $x_E = x_F + x_G$.

Définition 18.5

Soient E un $\mathbb{K}$ -espace vectoriel et F et G , deux sous-espaces vectoriels de E . F et G sont **supplémentaires** dans E si et seulement si $F + G = E$ et l'intersection de F et de G est réduite à l'élément neutre, $F \cap G = \{0_E\}$, c'est-à-dire si et seulement si $E = F \oplus G$.

Exemple

Reprenons l'exemple précédent qui illustre la somme directe, nous avons bien $F \cap G = \{0_E\}$. D'autre part, $F + G = \mathbb{R}^3$ donc F et G sont supplémentaires.

Application

En physique, une force est une action mécanique capable de créer une accélération, c'est-à-dire une modification de la vitesse d'un objet ou d'une partie d'un objet, ce qui induit un déplacement ou une déformation de l'objet. Un représentant du vecteur force est caractérisé par quatre éléments : la direction, le sens, la norme et le point d'application (endroit où la force s'exerce).

Nous considérons l'ensemble V de tous les vecteurs de l'espace $\mathbb{R}^3$. Cet espace est un modèle habituel des différentes forces en Physique. Montrer qu'il s'agit d'un espace vectoriel.

Tout d'abord, il existe :

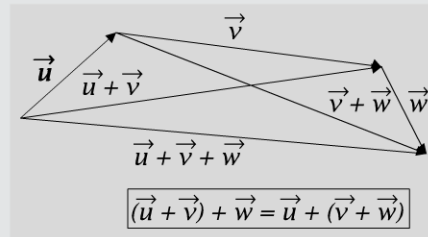
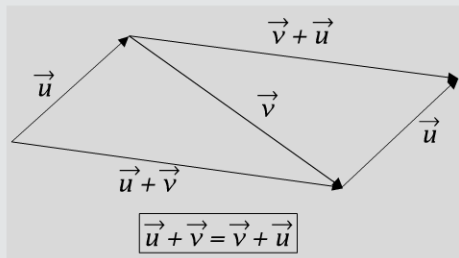
- une loi interne, l'addition, qui est définie par la règle du parallélogramme :

$$\forall \vec{u}, \vec{v} \in V, \vec{u} + \vec{v} \in V,$$

- une loi externe, la multiplication par un scalaire : $\forall \lambda \in \mathbb{R}, \forall \vec{u} \in V, \lambda \vec{u} \in V$ qui est défini comme le vecteur de longueur $|\lambda|$ fois plus grande que celle de $\vec{u}$, de même direction que $\vec{u}$ et de même sens que $\vec{u}$ si $\lambda > 0$ ou de sens contraire si $\lambda < 0$.

Alors, nous montrons de façon géométrique, en nous aidant des schémas ci-dessous :

- $\forall \vec{u}, \vec{v} \in V, \vec{u} + \vec{v} = \vec{v} + \vec{u}$,
- $\forall \vec{u}, \vec{v}, \vec{w} \in V, \vec{u} + (\vec{v} + \vec{w}) = (\vec{u} + \vec{v}) + \vec{w}$.



De plus,

- L'addition a un élément neutre puisque un segment de longueur nulle est un simple point et représente le vecteur nul, noté $\vec{0}$: $\forall \vec{u} \in V, \vec{u} + \vec{0} = \vec{0} + \vec{u} = \vec{u}$,
- $\forall \vec{u} \in V, \exists -\vec{u} \in V$ tel que $\vec{u} + -\vec{u} = \vec{0}$,
- $\forall \lambda, \mu \in \mathbb{R}, \forall \vec{u} \in V, \lambda(\mu \vec{u}) = (\lambda\mu) \vec{u}$,
- $\forall \lambda \in \mathbb{R}, \forall \vec{u}, \vec{v} \in V, \lambda(\vec{u} + \vec{v}) = \lambda \vec{u} + \lambda \vec{v}$,
- $\forall \lambda, \mu \in \mathbb{R}, \forall \vec{u} \in V, (\lambda + \mu) \vec{u} = \lambda \vec{u} + \mu \vec{u}$.

Nous en déduisons donc que V est un $\mathbb{R}$ -espace vectoriel.

Pour s'entraîner
(solutions p. 512)

18.1 Soit $E = \mathbb{R}^3$. Les sous-ensembles suivants sont-ils des sous-espaces vectoriels de E ?

1. $E_1 = \{\vec{u} = (x; y; z) \in \mathbb{R}^3; x + z = 0\}$,

2. $E_2 = \{\vec{u} = (x; y; z) \in \mathbb{R}^3; xz = 0\}$.

18.2 Soit E l'ensemble des fonctions définies de $\mathbb{R}$ dans $\mathbb{R}$. Montrer que l'ensemble $E_1 = \{f : \mathbb{R} \rightarrow \mathbb{R}; f \text{ est paire}\}$ est un sous-espace vectoriel de E .

18.3 Soit $E = \mathbb{R}[X]$. Montrer que l'ensemble des polynômes de degré inférieur ou égal à trois est un sous-espace vectoriel de E .

18.4 Soit $E = \{\vec{u} = (x; y) \in \mathbb{R}^2; x \geq 0, y \geq 0\}$, c'est-à-dire le premier quadrant du plan $(0; x; y)$.

1. Est-ce que $\vec{u} + \vec{v} \in E$ pour $\vec{u}, \vec{v} \in E$?

2. Identifier $\vec{u} \in E$ et $\lambda \in \mathbb{R}$ tels que $\lambda \vec{u} \notin E$. Cela suffit-il pour montrer que E n'est pas un $\mathbb{R}$ -espace vectoriel ?

18.5 Soit $F = \{\vec{u} = (x; y) \in \mathbb{R}^2; xy \geq 0\}$.

1. Est-ce que $\lambda \vec{u} \in F$ pour $\vec{u} \in F$ et $\lambda \in \mathbb{R}$?

2. Identifier $\vec{u}$ et $\vec{v} \in F$ tels que $\vec{u} + \vec{v} \notin F$. Cela suffit-il pour montrer que F n'est pas un $\mathbb{R}$ -espace vectoriel ?

Dans toute cette fiche, $\mathbb{K}$ désigne un corps commutatif et n un entier naturel non nul.

1. Familles génératrices, familles libres et bases

Définition 19.1

Soient I un ensemble, E un espace vectoriel sur le corps $\mathbb{K}$, $(\vec{u}_i)_{i \in I}$ une famille de vecteurs de E et $(\lambda_i)_{i \in I}$ une famille d'éléments de $\mathbb{K}$ presque tous nuls, c'est-à-dire tous nuls sauf éventuellement un nombre p fini d'entre eux. Nous appelons **combinaison linéaire** de vecteurs de la famille $(\vec{u}_i)_{i \in I}$ de coefficients $(\lambda_i)_{i \in I}$ la somme

$$\sum_{i \in I} \lambda_i \cdot \vec{u}_i = \lambda_{i_1} \cdot \vec{u}_{i_1} + \cdots + \lambda_{i_p} \cdot \vec{u}_{i_p} = \sum_{k=1}^p \lambda_{i_k} \cdot \vec{u}_{i_k},$$

où les i_k , $1 \leq k \leq p$, sont les seuls éléments de I pour lesquels $\lambda_{i_k} \neq 0$.

Définition 19.2

Soient I un ensemble, E un espace vectoriel sur le corps $\mathbb{K}$, $(\vec{u}_i)_{i \in I}$ une famille de vecteurs de E . Un vecteur $\vec{u}$ est une combinaison linéaire des vecteurs de la famille $(\vec{u}_i)_{i \in I}$ s'il existe une famille d'éléments de $\mathbb{K}$ presque tous nuls $(\lambda_i)_{i \in I}$ telle que u est la **combinaison linéaire** de vecteurs de la famille $(\vec{u}_i)_{i \in I}$ de coefficients $(\lambda_i)_{i \in I}$.

Définition 19.3

Soient I un ensemble, E un espace vectoriel sur le corps $\mathbb{K}$, $(\vec{u}_i)_{i \in I}$ une famille de vecteurs de E et F un sev de E .

- $(\vec{u}_i)_{i \in I}$ est une **famille génératrice de F** si tout élément de F s'écrit comme une combinaison linéaire des vecteurs de la famille $(\vec{u}_i)_{i \in I}$.
- Les éléments de $\mathbb{K}$, presque tous nuls, $(\lambda_i)_{i \in I}$ qui interviennent dans l'écriture de u comme combinaison linéaire des vecteurs de la famille $(\vec{u}_i)_{i \in I}$ sont des **coordonnées de $\vec{u}$ dans la famille $(\vec{u}_i)_{i \in I}$** .
- Si $(\vec{u}_i)_{i \in I}$ est une famille génératrice de F , alors nous disons que F est **engendré par $(\vec{u}_i)_{i \in I}$** .

Définition 19.4

Soient I un ensemble, E un espace vectoriel sur le corps $\mathbb{K}$, $(\vec{u}_i)_{i \in I}$ une famille de vecteurs de E . Nous appelons **sous-espace vectoriel engendré** par $(\vec{u}_i)_{i \in I}$ et nous notons $\text{Vect}((\vec{u}_i)_{i \in I})$, l'ensemble des combinaisons linéaires de ces vecteurs. Il s'agit du plus petit sous-espace vectoriel de E contenant ces vecteurs.

Définition 19.5

Les vecteurs $(\vec{u}_i)_{i \in I}$ forment une **famille libre** de E si et seulement si aucun des vecteurs de $(\vec{u}_i)_{i \in I}$ ne peut être obtenu comme combinaison linéaire des autres vecteurs de cette famille.

Définition 19.6

Si les vecteurs $(\vec{u}_i)_{i \in I}$ ne forment pas une famille libre, nous disons qu'ils sont **liés**.

Exemple

$\{(2; 0), (1; 3)\}$ est libre dans $\mathbb{R}^2$ puisque $(2; 0)$ ne s'exprime pas en fonction de $(1; 3)$.

Proposition 19.1

Avec les notations de la définition précédente, nous avons :

$$(\vec{v}_1; \dots; \vec{v}_p) \text{ famille libre} \iff \left\{ \sum_{i=1}^p \lambda_i \cdot \vec{v}_i = \vec{0} \Rightarrow \forall i \in \{1, \dots, p\}, \lambda_i = 0 \right\}.$$

Définition 19.7

Soit $(\vec{u}_i)_{i \in I}$ une famille d'éléments de F , un sev de E . Nous dirons que $(\vec{u}_i)_{i \in I}$ est une **base de F** si $(\vec{u}_i)_{i \in I}$ est à la fois une famille libre et génératrice de F .

Proposition 19.2

Soit F un sev d'un espace vectoriel E sur le corps $\mathbb{K}$. Soit $(\vec{u}_i)_{i \in I}$ une base de F . Tout vecteur $\vec{u} \in F$ s'écrit de **manière unique** comme combinaison linéaire des vecteurs de base $(\vec{u}_i)_{i \in I}$. En fait, nous avons :

$$(\vec{u}_i)_{i \in I} \text{ base de } F \iff \forall \vec{u} \in F, \exists! (\lambda_i)_{i \in I} \in \mathbb{K}^I, \text{ presque tous nuls, } \vec{u} = \sum_{k=1}^p \lambda_{i_k} \cdot \vec{u}_{i_k}.$$

Les scalaires $(\lambda_i)_{i \in I}$ sont les **coordonnées de $\vec{u}$ dans la base $(\vec{u}_i)_{i \in I}$** . Ces coordonnées sont uniques.

Définition 19.8

Un **espace vectoriel E** sur le corps $\mathbb{K}$ est de **dimension finie** s'il existe une base de E de cardinal fini.



Si E est un espace vectoriel de dimension finie sur le corps $\mathbb{K}$, alors nous avons :

$$(\vec{v}_1; \dots; \vec{v}_p) \text{ base de } E \iff \forall \vec{u} \in E, \exists! \lambda_1, \dots, \lambda_p \in \mathbb{K} / \vec{u} = \sum_{i=1}^p \lambda_i \cdot \vec{v}_i.$$

Exemples

$\{(1; 0); (0; 1)\}$ est une base de $\mathbb{R}^2$. $\{1; X; X^2; X^3\}$ est une base de $\mathbb{R}_3[X]$.

2. Dimension**Définition 19.9**

Un $\mathbb{K}$ -espace vectoriel E est dit de **dimension finie** si et seulement si E admet une famille génératrice finie. Dans ce cas, la **dimension** de E est le cardinal de toute base de E . Nous la notons $\dim(E)$.

Définition 19.10

Nous appelons **droite vectorielle** tout espace vectoriel ou sous-espace vectoriel de dimension une, **plan vectoriel** tout espace vectoriel ou sous-espace vectoriel de dimension deux et **hyperplan** tout sous-espace vectoriel de dimension $n - 1$.

Théorème 19.1

Soient E un $\mathbb{K}$ -espace vectoriel de dimension finie et $n = \dim(E)$. Alors :

- toutes les bases de E sont finies et ont le même cardinal qui est la dimension de E ;
- toute famille libre de E est finie et a au plus n éléments ;
- toute famille de E ayant au moins $n + 1$ éléments est donc liée ;
- toute famille génératrice de E a au moins n éléments ;

Théorème 19.2

Soient E un $\mathbb{K}$ -espace vectoriel de dimension finie avec $\dim(E) = n$ et $\mathcal{F}$ une famille finie d'éléments de E . Alors, les trois assertions suivantes sont équivalentes.

- $\mathcal{F}$ est une base.
- $\mathcal{F}$ est libre à n éléments.
- $\mathcal{F}$ est génératrice à n éléments.

Théorème 19.3

Soient E un $\mathbb{K}$ -espace vectoriel de dimension finie, F et G deux sous-espaces vectoriels de E .

- Tout sous-espace vectoriel de E est de dimension finie inférieure ou égale à $\dim(E)$.
- Si $F \subset G$ et $\dim(F) = \dim(G)$, alors $F = G$.
- Si F et G sont en somme directe alors $\dim(F \oplus G) = \dim(F) + \dim(G)$.

Théorème 19.4 Formule de Grassmann

Soit E un $\mathbb{K}$ -espace vectoriel de dimension finie. Alors, pour tous les sous-espaces vectoriels F, G de E , $\dim(F + G) = \dim(F) + \dim(G) - \dim(F \cap G)$.

Définition 19.11

Soient E un $\mathbb{K}$ -espace vectoriel et $\mathcal{F}$ une famille finie d'éléments de E . Nous appelons **rang** de $\mathcal{F}$, noté $\text{rg}(\mathcal{F})$, l'entier naturel $\dim(\text{Vect}(\mathcal{F}))$. C'est le plus grand cardinal des familles libres de $\mathcal{F}$.

Théorème 19.5

Soit E un espace vectoriel sur le corps $\mathbb{K}$ de dimension finie et F un sev de E .

- **Théorème de la base incomplète.** Toute famille libre de E peut être complétée pour former une base de E .
- De toute famille génératrice de E , nous pouvons extraire une base de E .

Soient les vecteurs $\vec{u}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}$, $\vec{u}_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 3 \end{pmatrix}$, $\vec{u}_3 = \begin{pmatrix} 2 \\ 1 \\ 1 \\ 1 \end{pmatrix}$, $\vec{u}_4 = \begin{pmatrix} -1 \\ 0 \\ -1 \\ 2 \end{pmatrix}$ et $\vec{u}_5 = \begin{pmatrix} 2 \\ 3 \\ 0 \\ 1 \end{pmatrix}$ appartenant à

$\mathbb{R}^4$. Soient également F un sous-espace de $\mathbb{R}^4$ engendré par $\{\vec{u}_1; \vec{u}_2; \vec{u}_3\}$ et G un sous-espace de $\mathbb{R}^4$ engendré par $\{\vec{u}_4; \vec{u}_5\}$. Donner la dimension de F et G .

Commençons par étudier F . Comme F est engendré par $\{\vec{u}_1; \vec{u}_2; \vec{u}_3\}$, $\dim(F) \leq 3$. Montrons d'autre part que la famille $\{\vec{u}_1; \vec{u}_2; \vec{u}_3\}$ est libre.



Partons de $\lambda_1 \vec{u}_1 + \lambda_2 \vec{u}_2 + \lambda_3 \vec{u}_3 = \vec{0}$ et montrons que cela entraîne nécessairement que $\lambda_1 = \lambda_2 = \lambda_3 = 0$.

En effet, si nous remplaçons dans cette équation $\vec{u}_1, \vec{u}_2$ et $\vec{u}_3$ par les valeurs données dans l'énoncé, cela nous ramène alors à résoudre les équations :

$$\lambda_1 + \lambda_2 + 2\lambda_3 = 0, \quad 2\lambda_1 + \lambda_2 + \lambda_3 = 0, \quad 3\lambda_1 + \lambda_2 + \lambda_3 = 0, \quad \lambda_1 + 3\lambda_2 + \lambda_3 = 0.$$

Si nous soustrayons la deuxième équation à la troisième, nous trouvons directement $\lambda_1 = 0$. À partir de la deuxième équation nous en déduisons que $\lambda_2 = -\lambda_3$. Si nous injectons ce résultat dans la première équation nous trouvons $\lambda_2 - 2\lambda_2 = -\lambda_2 = 0$ donc $\lambda_2 = 0$.

Nous avons bien montré que $\lambda_1 = \lambda_2 = \lambda_3 = 0$, la famille $\{\vec{u}_1; \vec{u}_2; \vec{u}_3\}$ est donc libre.

Nous en déduisons donc que $\dim(F) = 3$.

Concernant maintenant le sous-ensemble G , nous savons que G est engendré par $\{\vec{u}_4; \vec{u}_5\}$ donc $\dim(G) \leq 2$. De plus, $\vec{u}_4$ et $\vec{u}_5$ ne sont pas colinéaires (car $\vec{u}_4$ ne peut s'exprimer en fonction de $\vec{u}_5$) donc ils forment une famille libre.

Nous en déduisons que $\dim(G) = 2$.

Pour s'entraîner (solutions p. 512)

19.1 Parmi les familles suivantes de l'espace vectoriel $E = \mathbb{R}^3$, lesquelles forment des bases de E ?

- $E_1 = \{(1; 0; 1); (-1; 1; 2); (-2; 1; 2)\}$.
- $E_2 = \{(1; 0; 1); (-1; 1; 2); (0; 1; 3)\}$.
- $E_3 = \{(1; 0; 1); (-1; 1; 2)\}$.
- $E_4 = \{(1; 2; 3); (4; 0; -1); (3; 7; 9)\}$.

19.2 Dans $\mathbb{R}_3[X]$, soient les polynômes :

$$P_1(X) = X(X-1)(X-2), \quad P_2(X) = X(X-2)(X-3), \quad P_3(X) = X(X-1)(X-3) \text{ et } P_4(X) = (X-1)(X-2)(X-3).$$

- La famille $\{P_1; P_2; P_3; P_4\}$ est-elle libre ?
- Forme-t-elle une base de $\mathbb{R}_3[X]$?

19.3 Déterminer une base des sous-espaces vectoriels suivants.

- $E_1 = \{(x; y; z) \in \mathbb{R}^3 ; 2x - 5y + z = 0, y + z = 0\}$.
- $E_2 = \{(x; y; z; t) \in \mathbb{R}^4 ; x + z - 2t = 0, x + y + t = 0, 2x - y + 3z = 0\}$.
- $E_3 = \{(x_1; x_2; x_3; x_4) \in \mathbb{R}^4 ; x_1 = 2x_2 - x_3, x_4 = x_1 + x_2 + x_3\}$.

19.4 Nous considérons les sous-espaces vectoriels de $\mathbb{R}^4$, $F = \{(x; y; z; t) \in \mathbb{R}^4 ; x - 2y = 0\}$ et $G = \{(x; y; z; t) \in \mathbb{R}^4 ; y + 2z + 3t = 0\}$.

- Donner une base de F et sa dimension.
- Donner une base de G et sa dimension.
- Donner une base de $F \cap G$ et sa dimension.

Dans toute cette fiche, $\mathbb{K}$ désigne un corps commutatif.

1. Définitions et exemples

Définition 20.1

Soient E et F deux $\mathbb{K}$ -espaces vectoriels. Une **application linéaire** (ou **morphisme**) de E dans F est une application $f : E \longrightarrow F$ telle que

$$\forall (u; v) \in E \times E, f(u + v) = f(u) + f(v),$$

$$\forall u \in E, \forall \lambda \in \mathbb{K}, f(\lambda u) = \lambda f(u).$$

Nous notons par $\mathcal{L}(E, F)$ l'ensemble des applications linéaires de E dans F . $\mathcal{L}(E, F)$ est un $\mathbb{K}$ -espace vectoriel. Si E et F sont de dimension finie, $\dim(E) = n$ et $\dim(F) = p$, nous avons $\dim(\mathcal{L}(E, F)) = np$. Si de plus, f est bijective, nous disons que f est un **isomorphisme** de E sur F .



Une application linéaire f vérifie $f(0) = 0$.

Exemples

1. Soit E un $\mathbb{K}$ -espace vectoriel. Nous définissons l'**homothétie vectorielle** de rapport $\alpha \in \mathbb{K}^*$ par l'application $h_\alpha : E \longrightarrow E$ qui à $u \in E$ associe $h_\alpha(u) = \alpha u$. Cette application est linéaire. En effet, $\forall u, v \in E, \forall \lambda \in \mathbb{K}$,

$$h_\alpha(u + v) = \alpha(u + v) = \alpha u + \alpha v = h_\alpha(u) + h_\alpha(v),$$

$$h_\alpha(\lambda u) = \alpha(\lambda u) = \lambda(\alpha u).$$

2. Soit E un $\mathbb{K}$ -espace vectoriel. Nous définissons la **translation** f_t de vecteur $t \in E$ par l'application $f_t : E \longrightarrow E$ qui à $u \in E$, associe $f_t(u) = u + t$. Si t est non nul, alors la translation n'est pas une application linéaire. En effet,

$$f_t(0) = t \neq 0.$$

3. Les projecteurs (voir exercice 20.5) et les opérateurs de dérivation (voir l'Application de cette Fiche) sont également des applications linéaires.

Définition 20.2

Soient E un $\mathbb{K}$ -espace vectoriel et $f : E \longrightarrow E$ une application. L'application f est un **endomorphisme** de E si et seulement si f est linéaire. Nous notons $\mathcal{L}(E)$ l'ensemble des endomorphismes de E . Si de plus f est bijective, nous disons que f est un **automorphisme** de E . Nous notons par $\mathcal{GL}(E)$ l'ensemble des automorphismes de E .

Exemple

L'homothétie vectorielle est un automorphisme de E .

2. Propriétés**Théorème 20.1**

Soient E et F deux $\mathbb{K}$ -espaces vectoriels et $f : E \longrightarrow F$ une application. f est linéaire si et seulement si $\forall \lambda \in \mathbb{K}, \forall u, v \in E$,

$$f(\lambda u + v) = \lambda f(u) + f(v).$$

Théorème 20.2

Soient E et F deux $\mathbb{K}$ -espaces vectoriels, $n \in \mathbb{N}^*$ et $f : E \longrightarrow F$ une application linéaire. Alors, pour tous $\lambda_1, \dots, \lambda_n \in \mathbb{K}$, $x_1, \dots, x_n \in E$,

$$f\left(\sum_{i=1}^n \lambda_i x_i\right) = \sum_{i=1}^n \lambda_i f(x_i).$$

Théorème 20.3

Soient E, F, G des $\mathbb{K}$ -espaces vectoriels et $f : E \longrightarrow F$, $g : E \longrightarrow F$ et $h : F \longrightarrow G$ des applications linéaires.

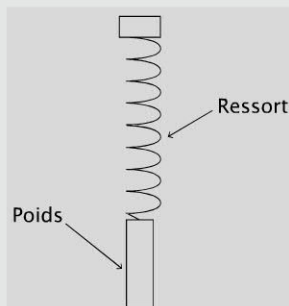
- La somme $f + g$ de deux applications linéaires, $f + g : E \longrightarrow F, u \mapsto f(u) + g(u)$, est une application linéaire.
- Le produit d'une application linéaire par un scalaire $\lambda \in \mathbb{K}$, soit $\lambda f : E \longrightarrow F, u \mapsto \lambda f(u)$ est une application linéaire.
- La composée de deux applications linéaires $h \circ g : E \longrightarrow G, u \mapsto (h(g(u))) = (h \circ g)(u)$ est une application linéaire.

Application

L'équation différentielle

$$y'' + \omega^2 y = 0$$

où ω est une constante réelle, régit toute une série de systèmes physiques comme par exemple l'oscillation d'un poids suspendu à un ressort, le mouvement d'un pendule ou la tension d'un circuit électrique LC (c'est-à-dire un circuit électrique contenant une bobine L et un condensateur C).



 Fiche 56

Définissons D sur $\mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$ et à valeurs dans $\mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$ où $\mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$ est l'ensemble des fonctions de $\mathbb{R}$ dans $\mathbb{R}$ de classe $\mathcal{C}^\infty$ par :

$$D : f \mapsto f'' + \omega^2 f.$$

Montrer que cette application D est une application linéaire.



Prenons pour cela f_1 et f_2 deux solutions de l'équation différentielle et $\lambda \in \mathbb{R}$. Nous voulons montrer que

$$D((f_1 + f_2)) = D(f_1) + D(f_2) \text{ et } D(\lambda f_1) = \lambda D(f_1).$$

Nous avons

$$\begin{aligned} D((f_1 + f_2)) &= (f_1 + f_2)'' + \omega^2(f_1 + f_2) \\ &= f_1'' + f_2'' + \omega^2 f_1 + \omega^2 f_2 \\ &= f_1'' + \omega^2 f_1 + f_2'' + \omega^2 f_2 \\ &= D(f_1) + D(f_2) \end{aligned}$$

$$\begin{aligned} D(\lambda f_1) &= \lambda f_1'' + \omega^2 \lambda f_1 \\ &= \lambda(f_1'' + \omega^2 f_1) \\ &= \lambda D(f_1). \end{aligned}$$

Donc D est bien une application linéaire.

Pour s'entraîner (solutions p. 513)

20.1 Montrer que la somme de deux applications linéaires de E dans F , deux $\mathbb{K}$ -espace vectoriels, est une application linéaire.

20.2 Montrer que le produit d'une application linéaire de E dans F , deux $\mathbb{K}$ -espace vectoriels, par un scalaire de $\mathbb{K}$ est une application linéaire.

20.3 Soient E , F et G trois $\mathbb{K}$ -espaces vectoriels, f une application linéaire de E dans F et g une application linéaire de F dans G . Montrer que la composée $g \circ f$ des deux applications linéaires f et g est une application linéaire de E dans G .

20.4 Montrer que les applications suivantes ne sont pas des applications linéaires.

1. $f : \mathbb{R}^2 \rightarrow \mathbb{R}, (x; y) \mapsto x + y + 1.$

2. $g : \mathbb{R}^2 \rightarrow \mathbb{R}, (x; y) \mapsto xy.$

20.5 Soient E un $\mathbb{K}$ -espace vectoriel et F et G deux sous-espaces vectoriels supplémentaires de E . Nous rappelons alors que si $x \in E$, $x = x_1 + x_2$ de façon unique avec $x_1 \in F$ et $x_2 \in G$.

1. Nous définissons le **projecteur** p de E dans E par l'application $p : x \mapsto p(x) = x_1$. C'est le projecteur sur F , parallèlement à G . Montrer que c'est une application linéaire.

2. Nous définissons la **symétrie** s_F de E dans E par l'application $s_F : x \mapsto s_F(x) = x_1 - x_2$. C'est la symétrie par rapport à F , parallèlement à G . Montrer que c'est une application linéaire.

20.6 Soit $\alpha \in \mathbb{K}^*$. L'homothétie de rapport α est définie par l'application h_α de E dans E telle que $x \mapsto h_\alpha(x) = \alpha x$. Montrer que c'est un automorphisme de E .

20.7 Soient r un scalaire et T l'application de $\mathbb{R}^2$ dans $\mathbb{R}^2$ qui à u associe $T(u) = ru$. T est appelée **contraction** lorsque $0 \leq r \leq 1$ et **dilatation** lorsque $r > 1$. Montrer que T est une application linéaire.

20.8

www



Soient E l'ensemble des polynômes de degré $\leq m$ et $f : E \rightarrow E$ définie par : $\forall P \in E$, $f(P) = P'$, où P' désigne la dérivée de P . Montrer que f est une application linéaire.

Noyau et image d'une application linéaire

1. Noyau et image d'une application linéaire

Définition 21.1

Soient E et F deux $\mathbb{K}$ -espaces vectoriels et $f : E \longrightarrow F$ une application linéaire. Nous appelons **noyau** de f et notons $\text{Ker}(f)$, l'ensemble des vecteurs de E dont l'image est le vecteur nul de F , autrement dit :

$$\text{Ker}(f) = f^{-1}(\{0_F\}) = \{u \in E ; f(u) = 0_F\}.$$

Exemples

1. Nous considérons l'application linéaire f de $\mathbb{R}^2$ dans $\mathbb{R}$ qui à tout couple $(x; y)$ associe x . Alors $\text{Ker}(f) = \{(x; y) \in \mathbb{R}^2 ; x = 0\} = \{(0; k) ; k \in \mathbb{R}\}$.
2. Soient W l'ensemble des fonctions continues sur $[a; b]$ et V l'ensemble des fonctions à valeurs réelles définies sur $[a; b]$ et de classe $\mathcal{C}^1$ sur $[a; b]$. Soit l'application linéaire $D : V \rightarrow W$ qui à toute fonction $f \in V$ associe sa dérivée première f' . Alors $\text{Ker}(D) = \{f \in V ; D(f) = 0_W\} = \{f \in V ; f' = 0\}$ et représente l'ensemble des fonctions constantes.

 Fiche 37

Théorème 21.1

Soient E et F deux $\mathbb{K}$ -espaces vectoriels. Alors, le noyau d'une application linéaire $f : E \longrightarrow F$ est un sous-espace vectoriel de E .

La notion de noyau induit le théorème suivant qui permet de montrer qu'une application linéaire est injective.

Théorème 21.2

Soient E et F deux $\mathbb{K}$ -espaces vectoriels et $f : E \longrightarrow F$ une application linéaire. Alors, nous avons l'équivalence suivante :

$$f \text{ injective} \iff \text{Ker}(f) = \{0_E\}.$$

Exemple

L'application linéaire f de $\mathbb{R}^2$ dans $\mathbb{R}$ qui à tout couple $(x; y)$ associe x n'est pas injective puisque nous avons montré dans l'exemple précédent que son noyau n'est pas réduit à l'élément neutre de $\mathbb{R}^2$.

Définition 21.2

Soient E et F deux $\mathbb{K}$ -espaces vectoriels et $f : E \longrightarrow F$ une application linéaire. Nous appelons **image** de f et notons $\text{Im}(f)$, l'ensemble des vecteurs de F qui ont un

antécédent dans E , autrement dit :

$$\text{Im}(f) = \{y \in F; \exists x \in E \text{ tel que } y = f(x)\}.$$

Théorème 21.3

Soient E et F deux $\mathbb{K}$ -espaces vectoriels. Alors, l'image d'une application linéaire $f : E \longrightarrow F$ est un sous-espace vectoriel de F .

Notons par ailleurs que l'image de l'application linéaire f est générée par l'image des éléments de n'importe quelle base de E . D'autre part, la notion d'image induit le théorème suivant qui permet de montrer qu'une application linéaire est surjective.

Théorème 21.4

Soient E et F deux $\mathbb{K}$ -espaces vectoriels et $f : E \longrightarrow F$ une application linéaire. Alors, nous avons l'équivalence suivante :

$$f \text{ surjective} \iff \text{Im}(f) = F.$$

Exemple

Soit l'application f de $\mathbb{R}^2$ dans $\mathbb{R}^2$ qui à $(x; y)$ associe $(y; x)$. Nous avons $\text{Im}(f) = \mathbb{R}^2$ donc f est surjective et $\text{Ker}(f) = \{0_{\mathbb{R}^2}\}$ donc f est également injective, et donc bijective.

2. Rang d'une application linéaire

Définition 21.3

Soient E et F deux $\mathbb{K}$ -espaces vectoriels et $f : E \longrightarrow F$ une application linéaire. Supposons que le sous-espace vectoriel $\text{Im}(f)$ est de dimension finie. Le **rang** de f , noté $\text{rg}(f)$, est égal à la dimension de $\text{Im}(f)$:

$$\text{rg}(f) = \dim(\text{Im}(f)).$$

Proposition 21.1

Soient E et F deux $\mathbb{K}$ -espaces vectoriels, F est de dimension finie, et $f : E \longrightarrow F$ une application linéaire.

$$f \text{ surjective} \iff \text{rg}(f) = \dim(F).$$

Théorème 21.5 Théorème du rang

Soient E et F deux $\mathbb{K}$ -espaces vectoriels, E de dimension finie, et $f : E \longrightarrow F$ une application linéaire. Alors,

$$\dim(E) = \dim(\text{Im}(f)) + \dim(\text{Ker}(f)) = \text{rg}(f) + \dim(\text{Ker}(f)).$$

Nous en déduisons :

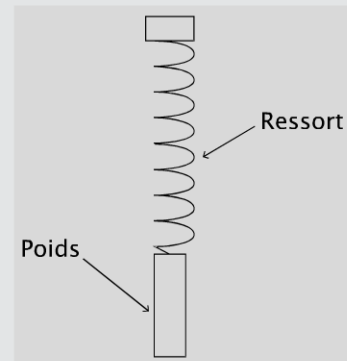
$$f \text{ injective} \iff \text{rg}(f) = \dim(E).$$

Application

L'équation différentielle

$$y'' + \omega^2 y = 0$$

où ω est une constante réelle régit toute une série de systèmes physiques comme par exemple l'oscillation d'un poids suspendu à un ressort, le mouvement d'un pendule et la tension d'un circuit électrique LC (c'est-à-dire contenant une bobine L et un condensateur C).



Notons $E = \mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$ l'espace vectoriel des fonctions de $\mathbb{R}$ dans $\mathbb{R}$ de classe $\mathcal{C}^\infty$ et D l'endomorphisme de E qui à $f \in E$ associe $D(f) = f'' + \omega^2 f$. L'ensemble des solutions de cette équation différentielle est le noyau de l'endomorphisme D . Déterminons ce noyau.

$$\begin{aligned} \text{Ker}(D) &= \{f \in E ; D(f) = 0\} \\ &= \{f \in E ; f'' + \omega^2 f = 0\}. \end{aligned}$$

D'après la théorie des équations différentielles, les solutions de l'équation différentielle de départ sont de la forme :

$$y(t) = c_1 \cos \omega t + c_2 \sin \omega t,$$

avec c_1 et c_2 des constantes réelles à définir en fonction des conditions expérimentales initiales.

Nous en déduisons que $\text{Ker}(D) = \{t \mapsto c_1 \cos \omega t + c_2 \sin \omega t ; c_1, c_2, \omega \in \mathbb{R}\}$.

 Fiche 37

 Fiche 56

Pour s'entraîner (solutions p. 514)

21.1 Les applications linéaires suivantes sont-elles injectives, surjectives, bijectives ?

1. $f_1 : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ définie par $f_1(x; y) = (2x; 0; x)$.

2. $f_2 : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ définie par $f_2(x; y) = (2y; x; 0)$.

3. $f_3 : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ définie par $f_3(x; y; z) = (2x - y + z; x - y - z; x + 2z)$.

4. $f_4 : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ définie par $f_4(x; y; z) = (x + y; x - z)$.

5. $f_5 : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ définie par $f_5(x; y) = (x + 2y; 3x - y; x + y)$.

21.2 Nous considérons $B = \{e_1; e_2; e_3\}$ une base de $\mathbb{R}^3$ et l'endomorphisme f de $\mathbb{R}^3$ défini par :

$$f(e_1) = e_1, f(e_2) = -e_1, f(e_3) = e_3.$$

1. Déterminer l'image par f d'un élément $u = xe_1 + ye_2 + ze_3$ de $\mathbb{R}^3$.

2. Déterminer le noyau et l'image de f et donner une base de chacun d'eux.

3. Montrer que $f \circ f = f$.

21.3 Soient E un $\mathbb{K}$ -espace vectoriel et F et G deux espaces vectoriels supplémentaires de E . Nous rappelons alors que si $x \in E$, $x = x_1 + x_2$ de façon unique avec $x_1 \in F$ et $x_2 \in G$.

1. Nous définissons le projecteur p de E dans E par l'application linéaire $p : x \mapsto p(x) = x_1$. C'est le projecteur sur F , parallèlement à G . Déterminer le noyau de p . L'application est-elle injective ?

2. Nous définissons la symétrie s_F de E dans E par l'application linéaire $s_F : x \mapsto s_F(x) = x_1 - x_2$. C'est la symétrie par rapport à F , parallèlement à G . Déterminer le noyau de s_F . L'application est-elle injective ?

21.4 Soient $E = \mathbb{R}_3[X]$ et $f : E \rightarrow E$ définie par :

$$\forall P \in E, f(P) = 2P - (X - 1)P'.$$

1. L'application linéaire f est-elle injective ?
2. L'application linéaire f est-elle surjective ?

21.5 Soient $k \in \mathbb{K}^*$. L'homothétie de rapport k est définie par l'application h_k de E dans E telle que $x \mapsto h_k(x) = kx$. Montrer que c'est un automorphisme de E . L'idée est d'utiliser le noyau et l'image de cette application linéaire pour montrer qu'elle est bijective.

21.6 

Soient E l'ensemble des polynômes de degré ≤ 2 et $f : E \rightarrow E$ définie par :

$$\forall P \in E, f(P) = P'.$$

Donner le noyau de f .

Dans toute cette Fiche, $\mathbb{K}$ désigne un corps.

1. Définitions et propriétés

Définition 22.1

Soient p et q deux entiers supérieurs ou égaux à 1. Une **matrice** A est un tableau d'éléments de $\mathbb{K}$ comportant p lignes et q colonnes. Nous notons par $(a_{ij})_{1 \leq i \leq p, 1 \leq j \leq q}$ les **coefficients** de la matrice A , a_{ij} étant le terme correspondant à la i -ème ligne et à la j -ème colonne de A . Dans la pratique, nous représenterons toujours A par un tableau rectangulaire de la forme :

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & \vdots & \dots & a_{1q} \\ a_{21} & a_{22} & \dots & \vdots & \dots & a_{2q} \\ \vdots & \ddots & \ddots & \vdots & \ddots & \vdots \\ \dots & \dots & \dots & a_{ij} & \dots & \dots \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{p1} & a_{p2} & \dots & \vdots & \dots & a_{pq} \end{pmatrix} = (a_{ij})_{1 \leq i \leq p, 1 \leq j \leq q} = (a_{ij})_{i,j}.$$



- Les matrices sont un outil pratique pour l'étude des applications linéaires.
- Nous disons que A est une matrice (p, q) à coefficients dans $\mathbb{K}$. (p, q) est la **taille** (ou le **format**) de la matrice. L'ensemble des matrices (p, q) à coefficients dans $\mathbb{K}$ sera noté $\mathcal{M}(p, q, \mathbb{K})$ dans la suite.
- Si $p = q$, alors A est une **matrice carrée**. L'ensemble des matrices carrées à p lignes et p colonnes à coefficients dans $\mathbb{K}$ sera noté $\mathcal{M}(p, \mathbb{K})$. Dans ce cas, les a_{ii} pour $i = 1, \dots, p$ sont les éléments diagonaux de A et $(a_{11}, \dots, a_{pp})$ est la **diagonale** de A . La **trace** de A , notée $tr(A)$ est la somme de ces éléments diagonaux, soit $tr(A) = \sum_{i=1}^p a_{ii}$.
- Si $p = 1$, alors A est une **matrice ligne**. Si $q = 1$, alors A est une **matrice colonne**.
- La **transposée** de $A \in \mathcal{M}(q, p, \mathbb{K})$, notée par ${}^tA \in \mathcal{M}(p, q, \mathbb{K})$ est la matrice dont les colonnes sont les lignes de A : ${}^tA = (a_{ji})_{1 \leq j \leq q, 1 \leq i \leq p}$.
- La **matrice nulle** de $\mathcal{M}(p, q, \mathbb{K})$ est la matrice $0_{(p,q)}$ dont tous les coefficients $(a_{ij})_{i,j}$ sont nuls.
- La **matrice identité** I_p est la matrice de $\mathcal{M}(p, \mathbb{K})$ qui n'a que des 1 sur la diagonale et des 0 partout ailleurs.

Définition 22.2 Somme de deux matrices et multiplication d'une matrice par un scalaire

- Soient $A = (a_{ij})_{i,j}$ et $B = (b_{ij})_{i,j}$ deux matrices de $\mathcal{M}(p, q, \mathbb{K})$. Alors la somme $A + B$ est la matrice de $\mathcal{M}(p, q, \mathbb{K})$ définie par $A + B = (a_{ij} + b_{ij})_{\{1 \leq i \leq p, 1 \leq j \leq q\}}$.
- Soient $\lambda \in \mathbb{K}$ et $A = (a_{ij})_{i,j} \in \mathcal{M}(p, q, \mathbb{K})$. Alors λA est une matrice de $\mathcal{M}(p, q, \mathbb{K})$ définie par $\lambda A = (\lambda a_{ij})_{\{1 \leq i \leq p, 1 \leq j \leq q\}}$.



- La multiplication par un scalaire est associative : soient $\lambda, \mu \in \mathbb{K}$ et $A \in \mathcal{M}(p, q, \mathbb{K})$, alors $(\lambda\mu)A = \lambda\mu A = \lambda(\mu A)$.
- Nous ne pouvons additionner que des matrices de même taille.
- $\mathcal{M}(p, q, \mathbb{K})$ muni des lois internes et externes précédemment énoncées est un $\mathbb{K}$ -espace vectoriel.

Définition 22.3 Produit de deux matrices

Soient $A = (a_{ij})_{i,j} \in \mathcal{M}(p, q, \mathbb{K})$ et $B = (b_{jk})_{j,k} \in \mathcal{M}(q, r, \mathbb{K})$ avec p, q et r des entiers non nuls. Alors le produit AB est une matrice C de $\mathcal{M}(p, r, \mathbb{K})$ dont les coefficients $(c_{ik})_{1 \leq i \leq p, 1 \leq k \leq r}$ vérifient $c_{ik} = \sum_{j=1}^q a_{ij}b_{jk}$.

Exemple

Soient $A = \begin{pmatrix} 2 & 1 \\ 1 & 3 \\ 0 & 1 \end{pmatrix}$ et $B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$. Alors $AB = \begin{pmatrix} 2 \times 1 + 1 \times 1 = 3 & 1 \\ 4 & 3 \\ 1 & 1 \end{pmatrix}$.



- Nous ne pouvons effectuer le produit des matrices A par B que si le nombre de colonnes de A est égal au nombre de lignes de B !
- Soient $A \in \mathcal{M}(p, q, \mathbb{K})$ et $B \in \mathcal{M}(q, r, \mathbb{K})$. Le produit AB est possible, mais pour effectuer le produit BA , il faudrait que $r = p$. Si c'est le cas, nous n'aurons pas nécessairement $AB = BA$. Le **produit matriciel n'est pas commutatif** !

Théorème 22.1

- Le produit matriciel est associatif : soient A, B, C des matrices de tailles respectives $(p, q), (q, r), (r, s)$ alors $A(BC) = (AB)C = ABC$.
- Le produit matriciel est distributif à droite par rapport à l'addition : soient $A \in \mathcal{M}(p, q, \mathbb{K})$ et $B, C \in \mathcal{M}(q, r, \mathbb{K})$, alors $A(B + C) = AB + AC$.
- Le produit matriciel est distributif à gauche par rapport à l'addition : soient $B, C \in \mathcal{M}(p, q, \mathbb{K})$ et $A \in \mathcal{M}(q, r, \mathbb{K})$, alors $(B + C)A = BA + CA$.

2. Matrices carrées

Définition 22.4

La matrice carrée $A \in \mathcal{M}(p, \mathbb{K})$ est **inversible** (ou **régulière**) s'il existe une matrice $B \in \mathcal{M}(p, \mathbb{K})$ telle que $AB = I_p = BA$. Si B existe, alors B est unique et $B = A^{-1}$.

Théorème 22.2

Soient A et B deux matrices inversibles de $\mathcal{M}(p, \mathbb{K})$. Alors AB est inversible et $(AB)^{-1} = B^{-1}A^{-1}$.

Définition 22.5

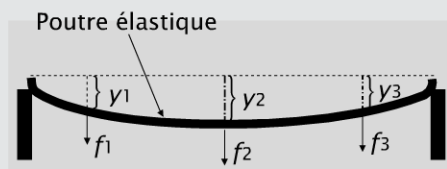
- Une matrice $A = (a_{ij})_{1 \leq i, j \leq p}$ de $\mathcal{M}(p, \mathbb{K})$ est **diagonale** si $a_{ij} = 0$ pour $i \neq j$.
- Une matrice $A = (a_{ij})_{1 \leq i, j \leq p}$ de $\mathcal{M}(p, \mathbb{K})$ est **triangulaire supérieure** (resp. **triangulaire inférieure**) si $a_{ij} = 0$ pour $i > j$ ($a_{ij} = 0$ pour $i < j$).
- Une matrice à la fois triangulaire supérieure et triangulaire inférieure est diagonale.
- Une matrice $A = (a_{ij})_{1 \leq i, j \leq p} \in \mathcal{M}(p, \mathbb{K})$ est **symétrique** (resp. **antisymétrique**) si $a_{ij} = a_{ji}$, $\forall i, j$ (resp. $a_{ij} = -a_{ji}$, $\forall i \neq j$).

Application

En Physique, intéressons-nous par exemple au fléchissement d'une poutre élastique. Considérons une poutre horizontale appuyée sur ses extrémités et soumise à des charges en différents points. Si $\vec{f}$ définit le vecteur des forces exercées en différents points de la poutre et $\vec{y}$ le vecteur des fléchissements (déplacements), alors, en vertu de la loi de Hooke,

$$\vec{y} = D\vec{f},$$

où D est la matrice dite de flexibilité. Trouver l'inverse D^{-1} de la matrice D , appelé matrice de rigidité et vérifiant $\vec{f} = D^{-1}\vec{y}$.



(D'après le livre *Algèbre linéaire : Théorie, exercices et application*, de David-C. Lay, Éditions De Boeck, 2004.)

Le calcul de l'inverse d'une matrice peut donc s'avérer très utile dans les applications. Cherchons par exemple l'inverse de la matrice $D = \begin{pmatrix} -3 & 5 & 6 \\ -1 & 2 & 2 \\ 1 & -1 & -1 \end{pmatrix}$.

Nous utilisons pour cela un algorithme qui consiste à travailler à partir du double tableau $[D \mid I]$ où I désigne la matrice identité de même taille que la matrice D . En effectuant diverses

opérations sur les lignes du double tableau, notre but est de faire apparaître la matrice I dans la partie gauche du double tableau, la partie droite donnant alors l'inverse recherché D^{-1} . Dans notre exemple, le double tableau peut s'écrire

$$\left(\begin{array}{cccccc|cccc} -3 & 5 & 6 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 2 & 2 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array} \right).$$

Nous inversons les lignes L_1 et L_3 , puis par les opérations $L_2 \leftarrow L_2 + L_1$, $L_3 \leftarrow L_3 + 3L_1$, puis les opérations $L_1 \leftarrow L_1 + L_2$, $L_3 \leftarrow L_3 - 2L_2$, nous obtenons successivement :

$$\left(\begin{array}{cccccc|cccc} 1 & -1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 3 & 1 & 0 & 3 & 0 & 0 & 1 & 1 \end{array} \right); \left(\begin{array}{cccccc|cccc} 1 & 0 & 0 & 0 & 1 & 2 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & -2 & 1 & 0 & 0 & 1 & 1 \end{array} \right).$$

Une dernière opération $L_2 \leftarrow L_2 - L_3$ permet d'obtenir :

$$\left(\begin{array}{cccccc|cccc} 1 & 0 & 0 & 0 & 1 & 2 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & -2 & 1 & 0 & 0 & 1 & 1 \end{array} \right).$$

Nous en déduisons que $D^{-1} = \begin{pmatrix} 0 & 1 & 2 \\ -1 & 3 & 0 \\ 1 & -2 & 1 \end{pmatrix}$.

Pour s'entraîner (solutions p. 514)

22.1 Combien de colonnes possède la matrice B si le produit AB est une matrice de taille $(4, 5)$?

22.2 Trouver deux matrices A et B telles que nous ayons à la fois $2A - 3B = \begin{pmatrix} 1 & 0 \\ 2 & 3 \end{pmatrix}$ et $A - 2B = \begin{pmatrix} 1 & 3 \\ 2 & -1 \end{pmatrix}$.

22.3 Soient $A = \begin{pmatrix} 2 & 3 & 0 & 1 \\ 1 & 5 & 1 & 0 \\ 0 & 2 & 3 & 2 \end{pmatrix}$ et $B = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 1 & 3 \end{pmatrix}$. Calculer si cela est possible $-2A$, $B - 2A$, $A + 2B$, AB et BA .

22.4 Soient $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}$ et $B = \begin{pmatrix} 1 & 3 & 5 & 7 \\ 2 & 4 & 6 & 8 \end{pmatrix}$.

Calculer si cela est possible AB et BA .

22.5 Soit $A = \begin{pmatrix} 0 & 1 & -1 \\ 4 & -3 & 4 \\ 3 & -3 & 4 \end{pmatrix}$. Calculer A^2 . En déduire A^{-1} .

22.6 Calculer l'inverse de $A = \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 2 & -3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$, $C = \begin{pmatrix} 2 & 2 & 3 \\ 1 & -1 & 0 \\ -1 & 2 & 1 \end{pmatrix}$.

Matrices et applications linéaires

Dans toute cette Fiche, $\mathbb{K}$ désigne un corps.

Dans certaines applications, il peut arriver qu'un problème décrit initialement dans une certaine base $\mathcal{B}$ trouve plus facilement une solution s'il est décrit dans une autre base $\mathcal{B}'$.

1. Matrice d'une application linéaire

Soient E un $\mathbb{K}$ -espace vectoriel de dimension finie n avec $n \in \mathbb{N}^*$ et $\mathcal{B} = \{e_1; \dots; e_n\}$ une base de E . Soit $x \in E$ avec $(a_1; \dots; a_n) \in \mathbb{K}^n$ les composantes de x dans $\mathcal{B}$. Alors, $x = \sum_{i=1}^n a_i e_i$.

 Fiche 19

Définition 23.1 Matrice d'une application linéaire

Soient E et F deux $\mathbb{K}$ -espaces vectoriels, de dimensions respectives q et p avec p, q des entiers non nuls. Considérons $\mathcal{B} = \{e_1; \dots; e_q\}$ une base de E et $\mathcal{B}' = \{e'_1; \dots; e'_p\}$ une base de F . Soit $f : E \rightarrow F$ une application linéaire.

Alors, la **matrice de f dans les bases $\mathcal{B}$ et $\mathcal{B}'$** , notée $Mat_{\mathcal{B}\mathcal{B}'} f$, est la matrice $A = Mat_{\mathcal{B}\mathcal{B}'} f$ de l'espace des matrices $\mathcal{M}(p, q, \mathbb{K})$ dont les coefficients de la j -ème colonne, pour $j = 1; \dots; q$, sont les coordonnées de $f(e_j)$ dans la base $\mathcal{B}'$. Plus précisément, $A = Mat_{\mathcal{B}\mathcal{B}'} f = (a_{ij})_{1 \leq i \leq p, 1 \leq j \leq q}$ avec $\forall 1 \leq j \leq q, f(e_j) = \sum_{i=1}^p a_{ij} e'_i$.

$$A = \begin{matrix} & \begin{matrix} f(e_1) & f(e_2) & \dots & f(e_q) \end{matrix} \\ \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1q} \\ a_{21} & a_{22} & \dots & a_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p1} & a_{p2} & \dots & a_{pq} \end{pmatrix} & \begin{matrix} \leftarrow e'_1 \\ \leftarrow e'_2 \\ \leftarrow \vdots \\ \leftarrow e'_p \end{matrix} \end{matrix}$$

Par exemple, le vecteur $f(e_1)$ s'écrit $f(e_1) = a_{11}e'_1 + a_{21}e'_2 + \dots + a_{p1}e'_p$.

Si $E = F$, f est un endomorphisme de E et $\mathcal{B} = \mathcal{B}'$, nous disons que A est la matrice de f dans la base $\mathcal{B}$, et nous notons $A = Mat_{\mathcal{B}} f$.

Théorème 23.1

Soient E et F deux $\mathbb{K}$ -espaces vectoriels de dimensions respectives q et p des entiers non nuls. Considérons $\mathcal{B}$ une base de E et $\mathcal{B}'$ une base de F . Soit $f : E \rightarrow F$ une application. Pour tout $u \in E$, notons X_u le vecteur des coordonnées de u dans la base $\mathcal{B}$ et Y_u le vecteur des coordonnées de $f(u)$ dans la base $\mathcal{B}'$.

- Si f est linéaire et si $A = Mat_{\mathcal{B}\mathcal{B}'} f$ alors $Y_u = AX_u$.
- Réciproquement, s'il existe une matrice $A \in \mathcal{M}(p, q, \mathbb{K})$ tel que pour tout $u \in E$, $Y_u = AX_u$ alors f est linéaire et $A = Mat_{\mathcal{B}\mathcal{B}'} f$.

2. Changements de bases

Définition 23.2

Soit E un $\mathbb{K}$ -espace vectoriel de dimension finie $n \in \mathbb{N}^*$. Soient $\mathcal{B} = \{e_1; \dots; e_n\}$ et $\mathcal{B}' = \{e'_1; \dots; e'_n\}$ deux bases de E . La **matrice de passage** de la base $\mathcal{B}$ à la base $\mathcal{B}'$, notée P est la matrice carrée dont les colonnes sont formées des coordonnées des vecteurs e'_i dans la base $\mathcal{B}$.

Exemple

Soit $E = \mathbb{R}^3$ muni de la base $\mathcal{B} = \{e_1; e_2; e_3\}$ avec $e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ et $e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. Soit $\mathcal{B}' = \{e'_1; e'_2; e'_3\}$ une base de $\mathbb{R}^3$ telle que $e'_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = e_1 + e_2$, $e'_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} = e_1 - e_2$ et $e'_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = e_1 + e_2 + e_3$. Alors,

$$P = \begin{array}{ccc} \begin{matrix} e'_1 \\ \downarrow \\ 1 \\ 1 \\ 0 \end{matrix} & \begin{matrix} e'_2 \\ \downarrow \\ 1 \\ -1 \\ 0 \end{matrix} & \begin{matrix} e'_3 \\ \downarrow \\ 1 \\ 1 \\ 1 \end{matrix} \\ \left(\begin{array}{ccc} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 0 & 0 & 1 \end{array} \right) & \begin{matrix} \leftarrow e_1 \\ \leftarrow e_2 \\ \leftarrow e_3 \end{matrix} \end{array}$$

Théorème 23.2

Soit P la matrice de passage de la base $\mathcal{B}$ à la base $\mathcal{B}'$. P est inversible et P^{-1} est la matrice de passage de la base $\mathcal{B}'$ à la base $\mathcal{B}$.

Théorème 23.3

- **Changement de base pour un vecteur.** Soit E un $\mathbb{K}$ -espace vectoriel de dimension $n \in \mathbb{N}^*$. Soient $\mathcal{B}$ et $\mathcal{B}'$ deux bases de E et P la matrice de passage de la base $\mathcal{B}$ à la base $\mathcal{B}'$. Soient $u \in E$ et X_u le vecteur des coordonnées de u dans la base $\mathcal{B}$ et X'_u le vecteur des coordonnées de u dans la base $\mathcal{B}'$. Alors

$$X_u = PX'_u \text{ ou } X'_u = P^{-1}X_u.$$

- **Changement de base pour une application linéaire.** Soient E et F deux $\mathbb{K}$ -espaces vectoriels de dimensions finies. Soient $\mathcal{B}_1$ et $\mathcal{B}_2$ deux bases de E et P la matrice de passage de $\mathcal{B}_1$ à $\mathcal{B}_2$. Soient $\mathcal{B}'_1$ et $\mathcal{B}'_2$ deux bases de F et Q la matrice de passage de $\mathcal{B}'_1$ à $\mathcal{B}'_2$. Enfin, soit $f : E \rightarrow F$, une application linéaire. Posons $A = \text{Mat}_{\mathcal{B}_1, \mathcal{B}'_1}(f)$ et $A' = \text{Mat}_{\mathcal{B}_2, \mathcal{B}'_2}(f)$. Alors,

$$A' = Q^{-1}AP.$$

Les matrices A et A' sont dites **équivalentes**. Elles représentent la même application linéaire dans des bases différentes.

Si $E = F$, f est un endomorphisme de E , posons $A = \text{Mat}_{\mathcal{B}_1}(f)$ et $A' = \text{Mat}_{\mathcal{B}_2}(f)$. Alors,

$$A' = P^{-1}AP.$$

Les matrices A et A' sont alors dites **semblables**.

Application Les chaînes de Markov

Supposons que nous étudions la population française qui réside en métropole. Chaque année, un vecteur tel que $\vec{x}_0 = \begin{pmatrix} 0,96 \\ 0,04 \end{pmatrix}$ indiquerait que 96 pourcents de la population française vit en métropole et que 4 pourcents de la population française vit à l'extérieur de la métropole (Dom-Tom et étranger). Il s'agit ici d'un **vecteur de probabilité** car la somme de ses composantes est égale à 1.

Une matrice carrée dont les colonnes sont des vecteurs de probabilité est une **matrice stochastique**. Une **chaîne de Markov** est une suite de vecteurs de probabilité associés à une matrice P telle que $\vec{x}_1 = P\vec{x}_0$, $\vec{x}_2 = P\vec{x}_1, \dots$. D'une manière générale, pour $k = 0; 1; 2; \dots$

$$\vec{x}_{k+1} = P\vec{x}_k.$$

Construisons la matrice P dite de « migration » entre la métropole et l'étranger, soit

$$P = \begin{pmatrix} 0,95 & 0,02 \\ 0,05 & 0,98 \end{pmatrix}.$$

Chaque année,

- 5 pourcents de la population de métropole part à l'étranger,
- 2 pourcents de la population de l'étranger rentre en métropole,
- 98 pourcent de la population de l'étranger reste à l'étranger,
- 95 pourcents de la population de métropole reste en métropole.

Au bout d'un an, quel est le pourcentage de la population française qui vit en métropole ? À l'étranger ?

(D'après le livre *Algèbre linéaire : Théorie, exercices et application*, de David-C. Lay, Éditions De Boeck, 2004.)

Au bout d'un an,

$$\vec{x}_1 = P\vec{x}_0 = \begin{pmatrix} 0,95 & 0,02 \\ 0,05 & 0,98 \end{pmatrix} \begin{pmatrix} 0,96 \\ 0,04 \end{pmatrix} = \begin{pmatrix} 0,91 \\ 0,09 \end{pmatrix}.$$

Nous en déduisons qu'au bout d'un an, 91 pourcents de la population française vit en métropole et 9 pourcents vit à l'étranger.

Pour s'entraîner
(solutions p. 515)

23.1 Soient quatre vecteurs de $\mathbb{R}^2$,
 $u_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, $u_2 = \begin{pmatrix} -4 \\ 3 \end{pmatrix}$, $v_1 = \begin{pmatrix} -5 \\ 4 \end{pmatrix}$,
 $v_2 = \begin{pmatrix} 2 \\ -7 \end{pmatrix}$ et soient $\mathcal{B} = \{u_1; u_2\}$ et
 $\mathcal{B}' = \{v_1; v_2\}$.

1. Montrer que $\mathcal{B}$ et $\mathcal{B}'$ sont deux bases de $\mathbb{R}^2$.

2. Soit le vecteur $u = \begin{pmatrix} -1 \\ 5 \end{pmatrix}$. Donner ses coordonnées dans la base $\mathcal{B}$ puis dans la base $\mathcal{B}'$.

3. Donner la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$.

4. Donner la matrice de passage de $\mathcal{B}'$ à $\mathcal{B}$.

23.2 Soient $\mathcal{B}$ et $\mathcal{B}'$ deux bases respectivement de $\mathbb{R}^4$ et $\mathbb{R}^3$ et f l'application linéaire dont la matrice dans les bases $\mathcal{B}$ et $\mathcal{B}'$ est A , avec

$$A = \begin{pmatrix} 1 & 3 & 1 & -2 \\ 2 & -2 & 5 & 1 \\ 0 & -1 & 4 & 4 \end{pmatrix}.$$

Donner l'image par f du vecteur dont les coordonnées dans $\mathcal{B}$ sont $\begin{pmatrix} -1 \\ 0 \\ 2 \\ 2 \end{pmatrix}$.

23.3 Soit $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ l'application linéaire définie par $f(x; y; z) = (x - y; x + 2y; 3x - 4z)$. Déterminer la matrice de f dans la base

$\{e_1; e_2; e_3\}$ de $\mathbb{R}^3$ définie par $e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ et $e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.

23.4 Soit $f : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ l'application linéaire définie par $f(x; y; z) = (x - y; x + y)$.

1. Déterminer la matrice de f dans les bases $\{e_1; e_2; e_3\}$ de $\mathbb{R}^3$ et $\{e'_1; e'_2\}$ de $\mathbb{R}^2$ définies par

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, e'_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ et } e'_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

2. Soit $\mathcal{B} = \{u_1; u_2\}$ une nouvelle base de $\mathbb{R}^2$ avec $u_1 = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$ et $u_2 = \begin{pmatrix} 4 \\ 0 \end{pmatrix}$. Déterminer la matrice de f dans la base $\{e_1; e_2; e_3\}$ pour l'espace de départ et dans $\mathcal{B}$ pour l'espace d'arrivée.

3. Soit $\mathcal{B}' = \{v_1; v_2; v_3\}$ une nouvelle base de $\mathbb{R}^3$ avec $v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$, $v_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ et $v_3 = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$. Déterminer la matrice de f dans la base $\mathcal{B}'$ pour l'espace de départ et dans la base $\{e'_1; e'_2\}$ pour l'espace d'arrivée.

4. Enfin, donner la matrice de f dans les bases $\mathcal{B}'$ pour l'espace de départ et $\mathcal{B}$ pour l'espace d'arrivée.

Dans toute cette Fiche, $\mathbb{K}$ désigne un corps commutatif et n un entier non nul. Le déterminant d'une matrice carrée est un outil mathématique très utile dans de nombreuses applications comme par exemple pour la résolution des systèmes linéaires. Nous présenterons quelques-unes d'entre elles dans une autre fiche. Nous définissons dans cette fiche les déterminants et leurs propriétés.

1. Définitions

Définition 24.1

- Soit $A = (a_{11})$ une matrice composée d'un seul élément. Alors, le déterminant de A , noté $\det(A)$ vaut a_{11} .
- Soit $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ une matrice carrée de taille 2. Alors, le déterminant de A vaut $\det(A) = a_{11}a_{22} - a_{12}a_{21}$.
- Soit $A = (a_{ij})_{1 \leq i, j \leq n}$ une matrice carrée de taille n . Soit M_{ij} la sous matrice obtenue en supprimant la i -ème ligne et la j -ème colonne de A . Le déterminant associé à la sous matrice M_{ij} d'ordre $n - 1$, noté $\det(M_{ij})$, est appelé **mineur** de l'élément a_{ij} de la matrice A . Nous définissons d'autre part le **cofacteur** A_{ij} de l'élément a_{ij} de la matrice A par la relation

$$A_{ij} = (-1)^{i+j} \det(M_{ij}).$$

Nous définissons enfin la **comatrice** de A (ou matrice des cofacteurs de A) par $\text{com}(A) = (A_{ij})_{1 \leq i, j \leq n}$.

Le déterminant de la matrice A est égal à la somme des produits obtenus en multipliant les éléments d'une ligne ou d'une colonne quelconque par les cofacteurs respectifs. Si nous développons par exemple le déterminant de A

- par rapport à la colonne j avec $1 \leq j \leq n$, alors $\det(A) = \sum_{i=1}^n a_{ij}A_{ij}$,
- par rapport à la ligne i avec $1 \leq i \leq n$, alors $\det(A) = \sum_{j=1}^n a_{ij}A_{ij}$.

Ces deux formules aboutissent bien sûr au même résultat, qui est le déterminant de A .

Exemple

Soit $A = \begin{pmatrix} -7 & -4 & 6 \\ 8 & 2 & 3 \\ 9 & -6 & 9 \end{pmatrix}$. Alors, si nous développons le déterminant de A par rapport à la première ligne, nous obtenons :

$$\det(A) = -7A_{11} - 4A_{12} + 6A_{13} = -7 \times 36 - 4 \times (-45) + 6 \times (-66) = -468.$$

Remarquons que les termes $(-1)^{i+j}$ dans la définition des cofacteurs A_{ij} peuvent être représentés par le tableau de signes $\begin{pmatrix} + & - & + \\ - & + & - \\ + & - & + \end{pmatrix}$.



De nombreuses notations sont utilisées afin de caractériser le déterminant. Les délimiteurs $| \ |$ sont réservés aux déterminants alors que les délimiteurs $[\]$ et $(\)$ sont associés aux matrices. D'une manière générale, si $A = (a_{ij})_{i,j} \in \mathcal{M}(n, \mathbb{K})$, alors les notations utilisées pour le déterminant sont :

$$\det(A) = |A| = \Delta = \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} = \det \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}.$$

2. Propriétés

Théorème 24.1

1. Si nous ajoutons à une colonne (ou à une ligne) de la matrice $A \in \mathcal{M}(n, \mathbb{K})$ une combinaison linéaire des autres colonnes (ou des autres lignes), nous obtenons une nouvelle matrice de même déterminant que A .
2. Si nous permutons deux colonnes (ou deux lignes) de la matrice $A \in \mathcal{M}(n, \mathbb{K})$, nous obtenons une nouvelle matrice dont le déterminant est $-\det(A)$.
3. Les déterminants d'une matrice $A \in \mathcal{M}(n, \mathbb{K})$ et de sa transposée tA sont égaux.
4. Si la matrice $A \in \mathcal{M}(n, \mathbb{K})$ comporte une ligne ou une colonne de 0 alors $\det(A) = 0$.
5. Si la matrice $A \in \mathcal{M}(n, \mathbb{K})$ a deux lignes ou deux colonnes proportionnelles (ou identiques) alors $\det(A) = 0$.
6. Si $A \in \mathcal{M}(n, \mathbb{K})$ est une matrice triangulaire supérieure ou inférieure, alors $\det(A)$ est égal au produit des éléments de sa diagonale.
7. Soit B la matrice obtenue à partir de la matrice $A \in \mathcal{M}(n, \mathbb{K})$ en multipliant une colonne ou une ligne par un scalaire $\lambda \in \mathbb{K}$, alors $\det(B) = \lambda \times \det(A)$.

Exemple

Soit $A = \begin{pmatrix} 1 & 4 & 2 \\ 2 & 5 & 4 \\ 3 & 6 & 6 \end{pmatrix}$. Alors $\det(A) = 0$ car les colonnes une et trois sont proportionnelles.

Théorème 24.2

Soit $A \in \mathcal{M}(n, \mathbb{K})$. Si $\det(A) \neq 0$, alors la matrice A est **inversible** c'est-à-dire qu'il existe une unique matrice $A^{-1} \in \mathcal{M}(n, \mathbb{K})$ telle que $AA^{-1} = A^{-1}A = I_n$, avec I_n la matrice identité de taille n .

Théorème 24.3

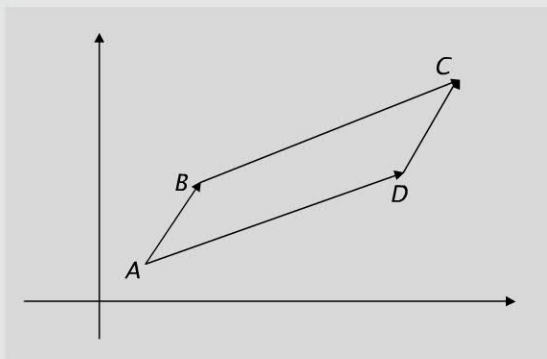
Soient A et B deux matrices appartenant à $\mathcal{M}(n, \mathbb{K})$. Alors $\det(AB) = \det(A) \times \det(B)$ et si A est inversible alors $\det(A) \neq 0$, $\det(A^{-1}) = \frac{1}{\det(A)}$.

Théorème 24.4

- Soit une matrice de taille $(2, 2)$, notée $A = \begin{pmatrix} u_1 & v_1 \\ u_2 & v_2 \end{pmatrix}$. Considérons les deux vecteurs du plan $\mathbb{R}^2$ de coordonnées données par les deux colonnes de la matrice : $\vec{u} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$, $\vec{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$. Si ces deux vecteurs sont non colinéaires, ils déterminent un parallélogramme. Alors, l'aire du parallélogramme est égale à la valeur absolue du déterminant de A .
- Soit une matrice de taille $(3, 3)$, notée $A = \begin{pmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{pmatrix}$. Considérons les trois vecteurs de l'espace $\mathbb{R}^3$ de coordonnées données par les trois colonnes de la matrice : $\vec{u} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}$, $\vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$ et $\vec{w} = \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix}$. Le produit mixte de ces trois vecteurs, noté $(\vec{u}, \vec{v}, \vec{w})$ est égal au déterminant $\begin{vmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{vmatrix}$. Si ces trois vecteurs sont non coplanaires, ils déterminent un parallélépipède. Alors, le volume du parallélépipède est égal à la valeur absolue du déterminant de A .

Application

Considérons les points A, B, C et D de coordonnées respectives dans le plan $\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 4 \end{pmatrix}, \begin{pmatrix} 8 \\ 6 \end{pmatrix}$ et $\begin{pmatrix} 6 \\ 3 \end{pmatrix}$. Calculer l'aire du parallélogramme $ABCD$.



Nous allons pour cela utiliser le dernier théorème donné dans la partie cours de cette fiche.

Dans cet exemple, les colonnes de ladite matrice sont déterminées par les coordonnées des vecteurs $\vec{AB}$ et $\vec{AD}$. La matrice a donc la forme $\begin{pmatrix} 2 & 5 \\ 3 & 2 \end{pmatrix}$.

Nous en déduisons que

$$\mathcal{A} = \left| \det \begin{pmatrix} 2 & 5 \\ 3 & 2 \end{pmatrix} \right| = |4 - 15| = 11.$$



Ici, les barres autour de $\det \begin{pmatrix} 2 & 5 \\ 3 & 2 \end{pmatrix}$ désignent la valeur absolue et non le déterminant !

Pour s'entraîner (solutions p. 515)

24.1 Montrer, sans le développer, que

$$\begin{vmatrix} 1+a & a & a \\ b & 1+b & b \\ c & c & 1+c \end{vmatrix} = 1 + a + b + c.$$

24.2 Montrer, avec le moins de calculs

$$\text{possibles, que } \begin{vmatrix} 73 & 7 & 3 \\ 10 & 1 & 0 \\ 51 & 5 & 1 \end{vmatrix} = 0.$$

24.3 Calculer les déterminants suivants en les développant selon les éléments de la première colonne. Calculer à nouveau les déterminants en les développant selon la deuxième ligne. Vérifier l'égalité des résultats.

$$1. \begin{vmatrix} 3 & 2 & 1 \\ 0 & 1 & 1 \\ 1 & 4 & 1 \end{vmatrix},$$

$$2. \begin{vmatrix} 2 & -1 & 1 \\ -1 & 6 & 1 \\ 0 & 2 & -1 \end{vmatrix}.$$

24.4 Calculer les deux déterminants suivants :

$$1. \begin{vmatrix} 1 & 2 & 4 \\ 0 & 1 & -2 \\ 5 & 6 & 2 \end{vmatrix}.$$

$$2. \begin{vmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \\ 4 & 5 & 6 & 7 \end{vmatrix}.$$

24.5 Calculer l'aire du parallélogramme dont les sommets sont données par les points de coordonnées : $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$, $\begin{pmatrix} -1 \\ 3 \end{pmatrix}$, $\begin{pmatrix} 3 \\ -2 \end{pmatrix}$ et $\begin{pmatrix} 4 \\ -5 \end{pmatrix}$.

24.6 Calculer le volume du parallélépipède dont un sommet est à l'origine et les autres sommets adjacents sont en les points

$$\begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix} \text{ et } \begin{pmatrix} 7 \\ 1 \\ 0 \end{pmatrix}.$$

Applications du calcul de déterminant

Dans toute cette Fiche, $\mathbb{K}$ désigne un corps et n un entier non nul.

Nous donnons dans cette Fiche deux applications du calcul de déterminant très utiles en pratique : le calcul de l'inverse d'une matrice et la résolution de systèmes linéaires par la méthode de Cramer.

1. Calcul pratique de l'inverse d'une matrice

Théorème 25.1

Soit $A = (a_{ij})_{1 \leq i, j \leq n}$ une matrice carrée appartenant à $\mathcal{M}(n, \mathbb{K})$. Si $\det(A) \neq 0$ alors A est inversible. Soit B la comatrice de A et tB sa transposée. Alors, l'inverse de A est donné par :

 Fiche 24

$$A^{-1} = \frac{1}{\det(A)} {}^tB.$$

Exemples

1. Soit $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ une matrice inversible. Alors $\det(A) = ad - bc \neq 0$, et

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

2. Soit $A = \begin{pmatrix} 1 & 3 & 3 \\ 2 & 5 & 3 \\ 1 & 2 & 1 \end{pmatrix}$. Alors $\det(A) = -1 \neq 0$ donc la matrice A est inversible. La comatrice

B de A est donnée par : $B = \begin{pmatrix} -1 & 1 & -1 \\ 3 & -2 & 1 \\ -6 & 3 & -1 \end{pmatrix}$. Nous obtenons donc en appliquant le théorème précédent

$$A^{-1} = \begin{pmatrix} 1 & -3 & 6 \\ -1 & 2 & -3 \\ 1 & -1 & 1 \end{pmatrix}.$$

2. Résolution de systèmes linéaires par la méthode de Cramer

Théorème 25.2

Soient une matrice $A = (a_{ij})_{1 \leq i, j \leq n} \in \mathcal{M}(n, \mathbb{K})$, un vecteur $b \in \mathbb{K}^n$, $b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$, et le système (S) suivant :

$$(S) : AX = b, \text{ d'inconnue } X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n.$$

Alors, les propriétés suivantes sont équivalentes :

- Le système (S) admet une solution et une seule.
- La matrice A est inversible.
- $\det(A) \neq 0$.

Si l'une de ces propriétés est vérifiée, le système (S) est appelé **système de Cramer**. De plus, si (S) est un système de Cramer, alors $\forall i = 1; \dots; n$,

$$x_i = \frac{1}{\det(A)} \times \begin{vmatrix} a_{11} & \dots & a_{1i-1} & y_1 & a_{1i+1} & \dots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \dots & a_{ni-1} & y_n & a_{ni+1} & \dots & a_{nn} \end{vmatrix}.$$

Exemple

Résoudre le système $(S) \begin{cases} 2x + y = 4 \\ 4x + 6y = 2 \end{cases}$.

Nous posons $X = \begin{pmatrix} x \\ y \end{pmatrix}$, $A = \begin{pmatrix} 2 & 1 \\ 4 & 6 \end{pmatrix}$ et $b = \begin{pmatrix} 4 \\ 2 \end{pmatrix}$.

(S) est alors équivalent à

$$(S) : AX = b.$$

Calculons $\det(A) = 2 \times 6 - 4 \times 1 = 8 \neq 0$. Le système (S) est donc un système de Cramer.

L'unique solution est donnée par :

$$x = \frac{1}{8} \begin{vmatrix} 4 & 1 \\ 2 & 6 \end{vmatrix} = \frac{22}{8} = \frac{11}{4} \text{ et } y = \frac{1}{8} \begin{vmatrix} 2 & 4 \\ 4 & 2 \end{vmatrix} = -\frac{3}{2}.$$

La solution est $X = \begin{pmatrix} \frac{11}{4} \\ -\frac{3}{2} \end{pmatrix}$.

Application

Considérons le système suivant dans lequel λ est un paramètre inconnu. Déterminer pour quelle(s) valeur(s) de λ le système admet une solution unique et décrire cette solution.

$$\begin{cases} \lambda x + y + z = 1 \\ x + \lambda y + z = \lambda \\ x + y + \lambda z = \lambda^2 \end{cases}$$



Le système est de la forme $AX = b$ avec $A = \begin{pmatrix} \lambda & 1 & 1 \\ 1 & \lambda & 1 \\ 1 & 1 & \lambda \end{pmatrix}$, $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ et $b = \begin{pmatrix} 1 \\ \lambda \\ \lambda^2 \end{pmatrix}$.

Pour calculer $\det(A)$, nous remplaçons d'abord la première colonne par la somme des trois colonnes de la matrice A . Cela fait apparaître un même coefficient qui est $\lambda + 2$ dans la première colonne que nous pouvons sortir à l'avant du calcul. Ensuite, nous remplaçons la colonne 2, par colonne 2-colonne 3, et ensuite nous remplaçons la colonne 3 par colonne 3-colonne 1. Enfin, pour terminer ce calcul de déterminant, il suffit de faire le produit des termes sur la diagonale de la matrice restante puisque celle-ci est triangulaire inférieure.

Ces étapes sont résumées ci-dessous :

$$\det(A) = \begin{vmatrix} \lambda & 1 & 1 \\ 1 & \lambda & 1 \\ 1 & 1 & \lambda \end{vmatrix} = \begin{vmatrix} \lambda+2 & 1 & 1 \\ \lambda+2 & \lambda & 1 \\ \lambda+2 & 1 & \lambda \end{vmatrix} = (\lambda+2) \begin{vmatrix} 1 & 0 & 0 \\ 1 & \lambda-1 & 0 \\ 1 & 1-\lambda & \lambda-1 \end{vmatrix} = (\lambda+2)(\lambda-1)^2.$$

Si $\lambda = -2$ ou $\lambda = 1$, $\det(A) = 0$ donc le système n'est pas de Cramer. Nous étudions les divers cas possibles.

1. Supposons que $\lambda \neq -2$ et $\lambda \neq 1$. Alors, le système admet une solution unique telle que :

$$x = \frac{1}{(\lambda+2)(\lambda-1)^2} \begin{vmatrix} 1 & 1 & 1 \\ \lambda & \lambda & 1 \\ \lambda^2 & 1 & \lambda \end{vmatrix} = \frac{1}{(\lambda+2)(\lambda-1)^2} \begin{vmatrix} 1 & 0 & 0 \\ \lambda & 0 & 1-\lambda \\ \lambda^2 & 1-\lambda^2 & \lambda-\lambda^2 \end{vmatrix} = \frac{(1-\lambda^2)(1-\lambda)}{(\lambda+2)(\lambda-1)^2} \begin{vmatrix} 1 & 0 & 0 \\ \lambda & 0 & 1 \\ \lambda^2 & 1 & \lambda \end{vmatrix}$$

$$= -\frac{(1+\lambda)}{(\lambda+2)}, \text{ obtenu en développant le déterminant selon la deuxième colonne,}$$

$$y = \frac{1}{(\lambda+2)(\lambda-1)^2} \begin{vmatrix} \lambda & 1 & 1 \\ 1 & \lambda & 1 \\ 1 & \lambda^2 & \lambda \end{vmatrix} = \frac{1}{(\lambda+2)(\lambda-1)^2} \begin{vmatrix} \lambda & 1 & 0 \\ 1 & \lambda & 1-\lambda \\ 1 & \lambda^2 & \lambda-\lambda^2 \end{vmatrix} = \frac{(1-\lambda)}{(\lambda+2)(\lambda-1)^2} \begin{vmatrix} \lambda & 1 & 0 \\ 1 & \lambda & 1 \\ 1 & \lambda^2 & \lambda \end{vmatrix}$$

$$= \frac{(1-\lambda)}{(\lambda+2)(\lambda-1)^2} \begin{vmatrix} \lambda & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & \lambda \end{vmatrix} = \frac{1}{\lambda+2}, \text{ en développant le déterminant selon la deuxième colonne,}$$

$$z = \frac{1}{(\lambda+2)(\lambda-1)^2} \begin{vmatrix} \lambda & 1 & 1 \\ 1 & \lambda & \lambda \\ 1 & 1 & \lambda^2 \end{vmatrix} = \frac{1}{(\lambda+2)(\lambda-1)^2} \begin{vmatrix} \lambda & 1 & 0 \\ 1 & \lambda & 0 \\ 1 & 1 & \lambda^2-1 \end{vmatrix} = \frac{(\lambda^2-1)}{(\lambda+2)(\lambda-1)^2} \begin{vmatrix} \lambda & 1 & 0 \\ 1 & \lambda & 0 \\ 1 & 1 & 1 \end{vmatrix}$$

$$= \frac{(\lambda+1)^2}{\lambda+2} \text{ en développant le déterminant selon la troisième colonne.}$$

2. Si $\lambda = 1$, le système se ramène à $x + y + z = 1$ et il y a une infinité de solutions de la forme $S = \{(x; y; z) \in \mathbb{R}^3; x = 1 - y - z\}$. L'ensemble des solutions est alors le plan passant par le point de coordonnées $\begin{pmatrix} 1/3 \\ 1/3 \\ 1/3 \end{pmatrix}$ et de normale $\vec{n} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

3. Si $\lambda = -2$, une résolution via la méthode du pivot nous permet d'arriver à une contradiction et à conclure qu'il n'y a pas de solution au système.

Pour s'entraîner
(solutions p. 516)

25.1 Inverser, si cela est possible, les matrices suivantes en utilisant la méthode de Cramer.

1. $A = \begin{pmatrix} 0 & 1 & -1 \\ 4 & -3 & 4 \\ 3 & -3 & 4 \end{pmatrix}$.

2. $B = \begin{pmatrix} 2 & 2 & 3 \\ 1 & -1 & 0 \\ -1 & 2 & 1 \end{pmatrix}$.

25.2 Soit $A = \begin{pmatrix} \lambda & 1 & 1 \\ 1 & \lambda & 1 \\ 1 & 1 & \lambda \end{pmatrix}$.

1. Pour quelle(s) valeur(s) du paramètre λ la matrice A est-elle inversible ?

2. Déterminer, lorsque cela est possible, l'inverse de la matrice A .

25.3 Résoudre par la méthode de Cramer les systèmes suivants :

1.
$$\begin{cases} -x + y - 2z = -4 \\ -x - y - z = -11 \\ 2x + y - z = 8 \end{cases}$$

2.
$$\begin{cases} 3x + 5y - z = 2 \\ 4z + 6 = 0 \\ 3x + 2y + z = 1 \end{cases}$$

25.4 Nous considérons le système linéaire suivant :

$$\begin{cases} mx + y + z + t = 1 \\ x + my + z + t = 0 \\ x + y + mz + t = 0 \\ x + y + z + mt = 0 \end{cases}$$

1. Pour quelle(s) valeur(s) de m ce système est-il un système de Cramer ?

2. Résoudre ce système.

Diagonalisation

Dans toute cette Fiche, $\mathbb{K}$ désigne un corps et n un entier non nul.

1. Valeurs propres, vecteurs propres, espaces propres

Définition 26.1

Soit $A = (a_{ij})_{1 \leq i, j \leq n} \in \mathcal{M}(n, \mathbb{K})$ une matrice carrée. Soit $\lambda \in \mathbb{K}$. Le scalaire λ est une **valeur propre** de A s'il existe un vecteur $u \in \mathbb{K}^n$ tel que $Au = \lambda u$. Le vecteur u est alors appelé **vecteur propre** de A **associé à la valeur propre** λ .

Définition 26.2

Le **polynôme caractéristique** P_A de $A \in \mathcal{M}(n, \mathbb{K})$ est défini par :

$$P_A(X) = \det(A - XI_n),$$

où I_n est la matrice identité de taille n .



- La somme des valeurs propres de la matrice A est égale à la trace de A .
- P_A est un polynôme de degré n .
- $P_A(0) = \det(A)$.
- La matrice A et sa transposée tA ont le même polynôme caractéristique.

Définition 26.3

Soient E un $\mathbb{K}$ -espace vectoriel, $\mathcal{B}$ une base de E et f un endomorphisme de E . Le polynôme caractéristique de la matrice de f dans la base $\mathcal{B}$ ne dépend pas de la base $\mathcal{B}$ choisie. Nous l'appelons **polynôme caractéristique de f** et le notons P_f .

Théorème 26.1 Théorème de Cayley-Hamilton

Soient E un $\mathbb{K}$ -espace vectoriel, f un endomorphisme de E et P_f son polynôme caractéristique. Alors, f annule P_f : $P_f(f) = 0$.

Exemple

Par exemple, si $P_f(X) = X^2 + 2X + 1$, alors $P_f(f) = 0$ signifie que l'application $f \circ f + 2f + 1Id_E$ est identiquement nulle sur E , où Id_E désigne l'application identité de E : $\forall u \in E, Id_E(u) = u$. Ainsi, $P_f(f) = 0$ signifie que $\forall u \in E, f \circ f(u) + 2f(u) + 1Id_E(u) = f(f(u)) + 2f(u) + u = 0$.

Théorème 26.2

Soit $A \in \mathcal{M}(n, \mathbb{K})$.

- Le scalaire $\lambda \in \mathbb{K}$ est une valeur propre de A si et seulement si $P_A(\lambda) = 0$.
- A admet au plus n valeurs propres.

Corollaire 26.1

Une matrice $A \in \mathcal{M}(n, \mathbb{K})$ est inversible si et seulement si 0 n'est pas une valeur propre de A .

Définition 26.4

Une valeur propre $\lambda \in \mathbb{K}$ de $A \in \mathcal{M}(n, \mathbb{K})$ est de multiplicité k si λ est une racine d'ordre k de P_A .



La somme des valeurs propres comptées avec multiplicité est égale à la trace de A .

**Point méthode**

Dans un premier temps, nous déterminons les valeurs propres de la matrice $A \in \mathcal{M}(n, \mathbb{K})$: ce sont les racines du polynôme caractéristique. Dans un second temps, nous cherchons les vecteurs propres de A . Les vecteurs propres de A associés à la valeur propre λ sont les vecteurs $u \in \mathbb{K}^n$ non nuls tels que :

$$Au = \lambda u.$$

Nous noterons par E_λ l'espace propre associé à la valeur propre λ c'est-à-dire l'ensemble des vecteurs propres associés à la valeur propre λ . Notons que cet espace propre est un sous-espace vectoriel.

2. Diagonalisation**Théorème 26.3**

Soient E un $\mathbb{K}$ -espace vectoriel, $\mathcal{B}$ une base de E , et une matrice $A \in \mathcal{M}(n, \mathbb{K})$. Supposons que le polynôme caractéristique de A , P_A , admet n racines distinctes $\lambda_1; \dots; \lambda_n$, c'est-à-dire que A admet n valeurs propres distinctes $\lambda_1; \dots; \lambda_n$.

Notons par u_i un vecteur propre de A associé à la valeur propre λ_i . Alors $\mathcal{B}' = \{u_1; \dots; u_n\}$ est une base de E . Notons enfin par P la matrice de passage de la base $\mathcal{B}$ à la base $\mathcal{B}'$ de E . La matrice A est dite **diagonalisable** et

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}.$$

La **transformation de similitude** est celle qui transforme A en $P^{-1}AP$.

Théorème 26.4

Soit $A \in \mathcal{M}(n, \mathbb{R})$ une matrice symétrique à coefficients réels, alors A est diagonalisable dans une base orthonormée formée de vecteurs propres.

Théorème 26.5

Soient $A \in \mathcal{M}(n, \mathbb{K})$ et $p \leq n$ un entier positif non nul. Supposons que le polynôme caractéristique P_A de A admet p racines distinctes $\lambda_1, \dots, \lambda_p$. Rappelons que $\lambda_1, \dots, \lambda_p$ sont aussi les valeurs propres de A . Notons m_i l'ordre de multiplicité de la valeur propre λ_i .

- La matrice A est diagonalisable si et seulement si pour tout $1 \leq i \leq p$ tel que $m_i \geq 2$, il existe m_i vecteurs propres linéairement indépendants associés à la valeur propre λ_i .
- Supposons A diagonalisable. Pour tout $1 \leq i \leq p$, notons

$$\{v_{\lambda_i,1}; \dots; v_{\lambda_i,m_i}\}$$

une famille libre de m_i vecteurs propres associés à la valeur propre λ_i . Alors la famille

$$\mathcal{B}' = \{v_{\lambda_1,1}; \dots; v_{\lambda_1,m_1}; v_{\lambda_2,1}; \dots; v_{\lambda_p,m_p}\}$$

est une base et A est diagonalisable dans cette base c'est-à-dire que la matrice $P^{-1}AP$ est diagonale et contient les valeurs propres de A rangées dans le même ordre que les vecteurs propres.



Dans le cas où la matrice étudiée n'est pas diagonalisable, il est souvent possible de la **triangulariser** c'est-à-dire de l'exprimer dans une base telle que la matrice $P^{-1}AP$ soit triangulaire supérieure.

Application

Cette application ne comporte volontairement pas de calculs, le but étant de montrer comment la diagonalisation peut s'avérer très pratique au niveau des applications.

Dans l'espace $\mathbb{R}^3$, muni de la base $\mathcal{B} = \{\vec{i}; \vec{j}; \vec{k}\}$ et du repère orthonormé $(O; \vec{i}; \vec{j}; \vec{k})$, soit un solide (S) positionné à l'origine O . Nous étudions la matrice d'inertie du solide (S) . La matrice d'inertie du solide (S) au point O est de la forme :

$$[I(S)] = \begin{pmatrix} A & -F & -E \\ -F & B & -D \\ -E & -D & C \end{pmatrix}$$

où A, B, C sont les moments d'inertie par rapport respectivement à $\vec{i}, \vec{j}, \vec{k}$:

$$A = \int \int_{(S)} (y^2 + z^2) dm, \quad B = \int \int_{(S)} (x^2 + z^2) dm, \quad C = \int \int_{(S)} (x^2 + y^2) dm$$

et D, E, F sont les produits d'inertie par rapport aux plans respectivement $(O; \vec{j}; \vec{k})$, $(O; \vec{i}; \vec{k})$ et $(O; \vec{i}; \vec{j})$:

$$D = \int \int_{(S)} (zy) dm, \quad E = \int \int_{(S)} (zx) dm, \quad F = \int \int_{(S)} (yx) dm.$$

Diagonaliser la matrice d'inertie $[I(S)]$.

La matrice d'inertie est toujours symétrique donc il existe toujours une base $\mathcal{B}'$ dans laquelle elle est diagonalisable. Cette base est dite « base propre d'inertie ». Les moments principaux A' , B' et C' sont les valeurs propres et la base du repère principal, soit $\mathcal{B}'$, correspond aux vecteurs propres associés. La matrice d'inertie est alors représentée dans la base $\mathcal{B}'$ par :

$$P^{-1}[I(S)]P = \begin{pmatrix} A' & 0 & 0 \\ 0 & B' & 0 \\ 0 & 0 & C' \end{pmatrix}$$

où P est la matrice de passage de la base $\mathcal{B}$ à la base $\mathcal{B}'$.

Pour s'entraîner (solutions p. 516)

26.1 Soit $A = \begin{pmatrix} 4 & 2 \\ -3 & -1 \end{pmatrix}$.

1. Montrer que 1 et 2 sont les valeurs propres de A .
2. Donner les vecteurs propres associés à chacune de ces valeurs propres.
3. Écrire A sous la forme PDP^{-1} où vous préciserez la forme des matrices P et D .

26.2 Soit $B = \begin{pmatrix} 2 & -3 \\ 3 & -4 \end{pmatrix}$.

1. Montrer que -1 est valeur propre double de B .
2. Donner le(s) vecteur(s) propre(s) associé(s) à cette valeur propre.
3. B est-elle diagonalisable ?

26.3 Soit $C = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$.

1. Montrer que $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ et $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ sont des vecteurs propres de C . À quelles valeurs propres sont-ils associés ?

2. C est-elle diagonalisable ?

26.4 Soit $E = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 1 \end{pmatrix}$.

1. Trouver les valeurs propres de E .
2. Déterminer les vecteurs propres associés à chacune de ces valeurs propres.
3. Écrire E sous la forme PDP^{-1} où vous préciserez la forme des matrices P et D .

26.5 Soit $F = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$.

1. Montrer que F a une valeur propre de multiplicité trois.
2. Donner le(s) vecteur(s) propre(s) associé(s) à cette valeur propre.
3. F est-elle diagonalisable ?

Applications de la diagonalisation

Dans toute cette Fiche, $\mathbb{K}$ désigne un corps et n un entier non nul.

Théorème 27.1 Calcul de la puissance d'une matrice

Soit p un entier positif non nul. Sous l'hypothèse qu'une matrice $A = (a_{ij})_{1 \leq i, j \leq n} \in \mathcal{M}(n, \mathbb{K})$ soit diagonalisable c'est-à-dire qu'il existe une matrice D diagonale et une matrice inversible P telles que $P^{-1}AP = D$, nous avons $A^p = PD^pP^{-1}$.

Exemple

Pour $p = 2$, nous retrouvons par exemple ce résultat facilement grâce à l'associativité du produit matriciel :

$$A^2 = (PDP^{-1})(PDP^{-1}) = PD(P^{-1}P)DP^{-1} = PD^2P^{-1}.$$

Une application essentielle de la diagonalisation est la résolution de systèmes différentiels linéaires .

Théorème 27.2

Soient un intervalle I de $\mathbb{R}$ et le système différentiel linéaire, défini pour $t \in I$ par :

$$(S) \begin{cases} x_1'(t) = a_{11}x_1(t) + a_{12}x_2(t) + \dots + a_{1n}x_n(t) + b_1(t) \\ x_2'(t) = a_{21}x_1(t) + a_{22}x_2(t) + \dots + a_{2n}x_n(t) + b_2(t) \\ \vdots = \vdots + \vdots + \vdots + \vdots + \vdots \\ x_n'(t) = a_{n1}x_1(t) + a_{n2}x_2(t) + \dots + a_{nn}x_n(t) + b_n(t) \end{cases}$$

où pour tout $i = 1; \dots; n$, les fonctions réelles $t \mapsto x_i(t)$ sont inconnues, les fonctions réelles $t \mapsto b_i(t)$ sont connues et les coefficients $(a_{ij})_{1 \leq i, j \leq n}$ sont des constantes réelles données. Pour résoudre un tel système, nous posons :

$$X(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{pmatrix}, \quad B(t) = \begin{pmatrix} b_1(t) \\ b_2(t) \\ \vdots \\ b_n(t) \end{pmatrix}, \quad A = \begin{pmatrix} a_{11} & \dots & \dots & a_{1n} \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ a_{n1} & \dots & \dots & a_{nn} \end{pmatrix}.$$

Alors, le système (S) peut s'écrire sous la forme matricielle $\frac{dX(t)}{dt} = AX(t) + B(t)$, pour $t \in I$.

Supposons que A soit diagonalisable, c'est-à-dire que nous puissions trouver une base telle que dans cette base, $D = P^{-1}AP$ soit diagonale, avec P la matrice de passage à cette nouvelle base. Alors, en posant, pour tout $t \in I$, les changements de variables $Y(t) = P^{-1}X(t)$, $C(t) = P^{-1}B(t)$, (S) peut se réécrire sous la forme

$$\frac{dY(t)}{dt} = DY(t) + C(t)$$

$$\text{où } Y(t) = \begin{pmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_n(t) \end{pmatrix}, C(t) = \begin{pmatrix} c_1(t) \\ c_2(t) \\ \vdots \\ c_n(t) \end{pmatrix}, \text{ c'est-à-dire}$$

$$(\tilde{S}) \begin{cases} y_1'(t) = \lambda_1 y_1(t) + c_1(t) \\ y_2'(t) = \lambda_2 y_2(t) + c_2(t) \\ \vdots = \vdots + \vdots \\ y_n'(t) = \lambda_n y_n(t) + c_n(t) \end{cases}.$$

La théorie classique sur les équations différentielles permet ensuite de résoudre aisément ce nouveau système. Nous en déduisons $Y(t)$, puis $X(t) = PY(t)$.

 Fiche 55

Définition 27.1

Soient $p \in \mathbb{N}^*$ et $n_0 \in \mathbb{N}$. Une **suite récurrente linéaire** d'ordre p est une suite $(U_n)_{n \geq n_0}$, à valeurs dans $\mathbb{K}$, définie par la relation de récurrence suivante : étant donnés p scalaires de $\mathbb{K}$, notés $a_0, a_1, \dots, a_{p-1}$ avec a_0 non nul, $\forall n \geq n_0$,

$$u_{n+p} = a_0 u_n + a_1 u_{n+1} + \dots + a_{p-1} u_{n+p-1} = \sum_{i=0}^{p-1} a_i u_{n+i}.$$

Exemple

1. Soient $q \in \mathbb{K}^*$ et une suite $(U_n)_{n \geq 0}$ à valeurs dans $\mathbb{K}$ vérifiant, $\forall n \geq 0$, la relation de récurrence $u_{n+1} = qu_n$. $(U_n)_{n \geq 0}$ est une suite récurrente linéaire d'ordre 1. Il s'agit ici d'une suite de type géométrique.
2. Une suite $(U_n)_{n \geq 0}$ à valeurs dans $\mathbb{K}$ vérifiant $u_{n+2} = au_{n+1} + bu_n$ où a et $b \in \mathbb{K}^*$ est une suite récurrente linéaire d'ordre 2.

Théorème 27.3

À chaque suite $(u_n)_{n \geq n_0}$, nous associons la suite $(X_n)_{n \geq n_0}$ telle que $X_n = \begin{pmatrix} u_n \\ u_{n+1} \\ \vdots \\ u_{n+p-1} \end{pmatrix}$. Pour tout $n \geq n_0$, la relation de récurrence définie à la Définition 27.1 sur la suite $(u_n)_{n \geq n_0}$ induit une relation de récurrence sur la suite $(X_n)_{n \geq n_0} : X_{n+1} = AX_n$, où

$$A = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & \dots & 1 \\ a_0 & a_1 & a_2 & \dots & a_{p-1} \end{pmatrix}.$$

Le terme général de la suite $(X_n)_{n \geq n_0}$ est alors déterminé par $X_n = A^n X_0$, ce qui permet de déduire le terme général de la suite $(u_n)_{n \geq n_0}$.

Application

Supposons que les tensions aux bornes de deux capacités d'un circuit électrique puissent être décrites par le système différentiel suivant :

$$\begin{pmatrix} v_1'(t) \\ v_2'(t) \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} v_1(t) \\ v_2(t) \end{pmatrix},$$

où $v_1(t)$ et $v_2(t)$ sont les tensions aux bornes des deux capacités au moment $t \in \mathbb{R}^+$. Supposons également que la tension initiale aux bornes de la première des deux capacités est de 4 volts tandis qu'aux bornes de la deuxième elle est de 2 volts. Trouver les expressions de $v_1(t)$ et $v_2(t)$ qui décrivent l'évolution de ces tensions dans le temps.

D'après l'énoncé, nous avons pour $t \in \mathbb{R}^+$,

$$\frac{dX(t)}{dt} = AX(t),$$

avec $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$, $X(t) = \begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix}$ et les conditions initiales $X(0) = \begin{pmatrix} 4 \\ 2 \end{pmatrix}$.

Les valeurs propres de A sont $\lambda_1 = 3$ et $\lambda_2 = -1$ auxquelles nous associons respectivement les vecteurs propres $u_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ et $u_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. La matrice A est donc diagonalisable et

$$D = P^{-1}AP,$$

avec $D = \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix}$ et $P = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$. Cela nous ramène donc à résoudre le système :

$$P^{-1} \frac{dX(t)}{dt} = DP^{-1}X(t),$$

ou encore, en posant $Y(t) = P^{-1}X(t)$,

$$\frac{dY(t)}{dt} = DY(t).$$



Le problème initial se ramène donc à résoudre les équations différentielles suivantes :

$$y_1'(t) = 3y_1(t) \text{ et } y_2'(t) = -y_2(t),$$

qui, par la théorie sur la résolution d'équations différentielles d'ordre 1, homogènes et à coefficients constants ont pour solution :

$$\forall t \in \mathbb{R}, y_1(t) = \lambda_1 e^{3t} \text{ et } y_2(t) = \lambda_2 e^{-t},$$

avec λ_1 et λ_2 des constantes réelles.

Revenons maintenant à la solution qui nous intéresse à savoir $v_1(t)$ et $v_2(t)$,

$$X(t) = PY(t) = \begin{pmatrix} y_1(t) + y_2(t) \\ y_1(t) - y_2(t) \end{pmatrix} = \begin{pmatrix} \lambda_1 e^{3t} + \lambda_2 e^{-t} \\ \lambda_1 e^{3t} - \lambda_2 e^{-t} \end{pmatrix}.$$

Nous utilisons enfin les conditions initiales pour déterminer λ_1 et λ_2 . Il suffit de résoudre $v_1(0) = \lambda_1 + \lambda_2 = 4$ et $v_2(0) = \lambda_1 - \lambda_2 = 2$. Nous trouvons $\lambda_1 = 3$ et $\lambda_2 = 1$ d'où la solution finale du problème :

$$X(t) = \begin{pmatrix} v_1(t) \\ v_2(t) \end{pmatrix} = \begin{pmatrix} 3e^{3t} + e^{-t} \\ 3e^{3t} - e^{-t} \end{pmatrix}.$$



Fiche 55

Pour s'entraîner (solutions p. 517)

27.1 Montrer que si une matrice $A = (a_{ij})_{1 \leq i, j \leq n} \in \mathcal{M}(n, \mathbb{K})$ est diagonalisable c'est-à-dire qu'il existe une matrice D diagonale et une matrice P inversible telles que $P^{-1}AP = D$, alors pour tout $p \in \mathbb{N}$, nous avons $A^p = PD^pP^{-1}$.

Astuce : utiliser un raisonnement par récurrence.

27.2 Soit $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$. Calculer A^n pour tout $n \in \mathbb{N}^*$.

27.3 Soit $B = \begin{pmatrix} 1 & 1 & 2 \\ 2 & 1 & 1 \\ 0 & 1 & 3 \end{pmatrix}$. Calculer B^n pour tout $n \in \mathbb{N}^*$.

27.4 Résoudre le système différentiel suivant, défini pour tout $t \in \mathbb{R}$ par :

$$(S) : \begin{cases} x_1'(t) = 2x_1(t) + x_2(t) - x_3(t) + 1 \\ x_2'(t) = x_1(t) + 2x_2(t) - x_3(t) + t \\ x_3'(t) = x_1(t) + x_2(t) + t^2 \end{cases}$$

27.5 Résoudre le système différentiel suivant, défini pour tout $t \in \mathbb{R}$ par :

$$(S) : \begin{cases} x_1'(t) = x_2(t) + e^t \\ x_2'(t) = -2x_1(t) + 3x_2(t) \end{cases}$$

avec les conditions initiales $x_1(0) = x_2(0) = 0$.

27.6 Trouver, pour $n \in \mathbb{N}^*$, le terme général de la suite réelle u_n définie par :

$$u_{n+2} = u_{n+1} + 2u_n,$$

avec $u_0 = 1$ et $u_1 = 0$.

Dans toute cette Fiche, $\mathbb{K}$ désigne un corps.

1. Formes bilinéaires, formes quadratiques

Définition 28.1

Une **forme bilinéaire** sur un $\mathbb{K}$ -espace vectoriel E est une application f de $E \times E$ dans $\mathbb{K}$ qui est linéaire par rapport à chaque variable : $\forall \lambda, \mu \in \mathbb{K}, \forall u, v, w, x \in E$,

$$f(\lambda u + v; \mu w + x) = \lambda \mu f(u; w) + \lambda f(u; x) + \mu f(v; w) + f(v; x).$$

Définition 28.2

Une forme bilinéaire f de $E \times E$ dans $\mathbb{K}$ est **symétrique** si $\forall (x; y) \in E \times E, f(x; y) = f(y; x)$.

Exemple

Soit $f : \mathbb{R}^2 \times \mathbb{R}^2 \longrightarrow \mathbb{R}$ définie par : $\forall x = (x_1; x_2) \in \mathbb{R}^2, \forall y = (y_1; y_2) \in \mathbb{R}^2, f(x; y) = x_1 y_2 + x_2 y_1$. L'application f est une forme bilinéaire symétrique.

Définition 28.3

Soit E un $\mathbb{K}$ -espace vectoriel. Une application q de E dans $\mathbb{K}$ est une **forme quadratique** sur E s'il existe une forme bilinéaire symétrique f telle que $\forall x \in E, q(x) = f(x; x)$. Cette forme bilinéaire symétrique f est unique. Elle est donnée par $\forall (x; y) \in E \times E$,

$$f(x; y) = \frac{1}{2} (q(x + y) - q(x) - q(y)).$$

f est appelée la **forme polaire** de q .



Dans un $\mathbb{K}$ -espace vectoriel E de dimension finie n , une forme quadratique q s'identifie facilement : c'est une expression polynomiale homogène en les coordonnées des éléments. En effet, il existe des scalaires $(a_{i,j})_{1 \leq i \leq j \leq n}$ tels que, pour tous $x \in \mathbb{K}^n$:

$$q(x) = q(x_1, \dots, x_n) = \sum_{i=1}^n a_{i,i} x_i^2 + 2 \sum_{1 \leq i < j \leq n} a_{i,j} x_i x_j.$$

Définition 28.4

Soient E , un $\mathbb{R}$ -espace vectoriel, q une forme quadratique sur E et f sa forme polaire.

- La forme quadratique q et sa forme polaire f sont dites **positives** si $\forall x \in E$,

$$q(x) \geq 0.$$

- La forme quadratique q et sa forme polaire f sont dites **définies positives** si $\forall x \in E$, x non nul,

$$q(x) > 0.$$

Exemple

Soit $f : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ définie par : $\forall x = (x_1; x_2, \dots; x_n) \in \mathbb{R}^n, \forall y = (y_1; y_2, \dots; y_n) \in \mathbb{R}^n$, $f(x; y) = \sum_{i=1}^n x_i y_i$. Alors $q(x) = \sum_{i=1}^n x_i^2$ est la forme quadratique associée à f et q est définie positive.

2. Espaces préhilbertiens réels

Définition 28.5

Soit E un $\mathbb{R}$ -espace vectoriel. Un **produit scalaire** sur E est une forme bilinéaire ϕ , symétrique, définie positive. (E, ϕ) est un **espace préhilbertien** réel. $\phi(x; y)$ se note $\langle x|y \rangle$ ou $x \cdot y$ ou encore $(x|y)$.

La norme sur E , associée au produit scalaire ϕ , est définie par :

$$\forall x \in E, \|x\| = \sqrt{\langle x|x \rangle}.$$

Théorème 28.1

Soit E un espace préhilbertien.

- **Égalité de polarisation.** $\forall x \in E, \forall y \in E$, nous avons

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2 + 2 \langle x|y \rangle.$$

- **Identité du parallélogramme.** $\forall x \in E, \forall y \in E$,

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2).$$

- **Inégalité de Cauchy-Schwarz.** $\forall x \in E, \forall y \in E$,

$$|\langle x|y \rangle| \leq \|x\| \times \|y\|.$$



L'inégalité de Cauchy-Schwarz est une égalité si et seulement si $\exists \lambda \in \mathbb{R}$ tel que $x = \lambda y$.

3. Espaces euclidiens

Définition 28.6

Un **espace vectoriel euclidien** ou **espace euclidien** E est un espace préhilbertien réel de dimension finie $n \in \mathbb{N}^*$.

Théorème 28.2

Soit E un espace euclidien. Il existe alors une base orthonormale $\mathcal{B} = \{e_1; \dots; e_n\}$ de E telle que $\forall x \in E, \forall y \in E$,

$$x = \sum_{i=1}^n \langle e_i | x \rangle e_i, \quad \|x\| = \sqrt{\sum_{i=1}^n x_i^2},$$

$$\langle x | y \rangle = {}^tXY,$$

avec $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ les matrices colonnes des composantes de x et y dans $\mathcal{B}$.

Application

1. Soit $E = \mathbb{R}^3$. Considérons l'application $f : E \times E \longrightarrow \mathbb{R}$ définie par $\forall x = (x_1; x_2; x_3) \in \mathbb{R}^3$, $\forall y = (y_1; y_2; y_3) \in \mathbb{R}^3$,

$$f(x; y) = 2x_1y_1 - 3x_3y_3.$$

La forme f est-elle bilinéaire ? symétrique ? positive ? définie positive ?

2. Soit q la forme quadratique sur $\mathbb{R}^3$ définie par : $\forall (x; y; z) \in \mathbb{R}^3$,

$$q(x; y; z) = x^2 - 2y^2 - xy - zx - 2yz.$$

Déterminer la forme bilinéaire f associée à q .

1.

- Tout d'abord, elle est bilinéaire. En effet, si nous prenons $\lambda, \mu \in \mathbb{R}, x = (x_1; x_2; x_3), y = (y_1; y_2; y_3), z = (z_1; z_2; z_3), t = (t_1; t_2; t_3) \in \mathbb{R}^3$,

$$\begin{aligned} f(\lambda x + y; \mu z + t) &= 2(\lambda x_1 + y_1)(\mu z_1 + t_1) - 3(\lambda x_3 + y_3)(\mu z_3 + t_3) \\ &= \lambda\mu(2x_1z_1 - 3x_3z_3) + 2\lambda x_1t_1 + 2\mu y_1z_1 + 2y_1t_1 \\ &\quad - 3\lambda x_3t_3 - 3\mu y_3z_3 - 3y_3t_3 \\ &= \lambda\mu f(x; z) + f(y; t) + \lambda f(x; t) + \mu f(y; z). \end{aligned}$$

- Montrons à présent que f est symétrique. Nous avons,

$$\begin{aligned} f(y; x) &= 2y_1x_1 - 3y_3x_3, \\ &= 2x_1y_1 - 3x_3y_3, \\ &= f(x; y). \end{aligned}$$

- Ensuite, f n'est pas positive car si nous choisissons par exemple le vecteur $x = (0, 0, 1)$,

$$f(x; x) = -3 < 0.$$

- Enfin, f n'est pas définie positive car il existe un vecteur $x \neq 0$ tel que $f(x; x) = 0$. Si nous prenons par exemple $x = (\sqrt{3}; 0; \sqrt{2})$, nous obtenons

$$f(x; x) = 2 \times 3 - 3 \times 2 = 0.$$

2. Posons $u = (x; y; z)$ et $v = (x'; y'; z') \in \mathbb{R}^3$. Alors, par définition,

$$\begin{aligned} f(u; v) &= \frac{1}{2} (q(u+v) - q(u) - q(v)), \\ &= xx' - 2yy' - \frac{1}{2}xy' - \frac{1}{2}x'y - \frac{1}{2}zx' - \frac{1}{2}z'x - yz' - y'z. \end{aligned}$$

Pour s'entraîner (solutions p. 518)

28.1 Soit $f : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}$ définie par :
 $\forall x = (x_1; x_2) \in \mathbb{R}^2, \forall y = (y_1; y_2) \in \mathbb{R}^2,$

$$f(x; y) = x_1y_2 + x_2y_1.$$

Montrer que f est une forme bilinéaire symétrique.

28.2 Soit $f : \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}$ définie par :
 $\forall x = (x_1; x_2; x_3) \in \mathbb{R}^3, \forall y = (y_1; y_2; y_3) \in \mathbb{R}^3,$

$$f(x; y) = x_1y_1^2 + x_2y_2^2 + x_3y_3^2.$$

Montrer que f n'est pas une forme bilinéaire.

28.3 Soit $E = C^0([0; 1])$ le $\mathbb{R}$ -espace vectoriel des fonctions continues sur $[0; 1]$. Pour $F, G \in E$, nous définissons l'application f par :

$$f(F; G) = \int_0^1 F(t)G(t)dt.$$

1. Montrer que f est une application bilinéaire symétrique définie positive.

2. Donner la forme quadratique q associée à f .

28.4 Soient $E = \mathbb{R}_2[X]$ et $P, Q \in E$ tels que :

$$\langle P|Q \rangle = P(-1)Q(-1) + P(0)Q(0) + P(1)Q(1).$$

Montrer que $\langle P|Q \rangle$ est un produit scalaire sur E .

28.5 

Soit $A = (a_{ij})_{1 \leq i, j \leq n}$ une matrice carrée d'ordre n à coefficients dans $\mathbb{R}$. Nous rappelons que la trace de A est définie par la somme de ses éléments diagonaux, soit $tr(A) = \sum_{i=1}^n a_{ii}$.

1. Montrer que $tr(A) = tr({}^tA)$.

2. Montrer que $\forall A, B \in \mathcal{M}(n, \mathbb{R}), tr(AB) = tr(BA)$.

3. Montrer que tr est une forme linéaire sur $\mathcal{M}(n, \mathbb{R})$.

4. Soit $\phi : \mathcal{M}(n, \mathbb{R}) \times \mathcal{M}(n, \mathbb{R}) \rightarrow \mathbb{R}$ qui à (M, N) associe $\phi(M, N) = tr({}^tMN)$. Montrer que ϕ est un produit scalaire sur $\mathcal{M}(n, \mathbb{R})$.

Orthogonalité, groupe orthogonal

Soit E un espace préhilbertien, muni du produit scalaire $\langle \cdot | \cdot \rangle$ et de la norme associée $\| \cdot \|$.

1. Généralités

Définition 29.1

- Deux vecteurs $x, y \in E$ sont **orthogonaux** si $\langle x | y \rangle = 0$. Nous le notons généralement par $x \perp y$.
- Une famille de vecteurs $(x_i)_{i \in I}$ de E est orthogonale si ses vecteurs sont deux à deux orthogonaux.
- Une famille de vecteurs $(x_i)_{i \in I}$ de E est **orthonormale** si elle est orthogonale et si les vecteurs qui la composent sont de norme égale à 1.

Théorème 29.1

- Une famille orthogonale de vecteurs non nuls est libre.
- **Théorème de Pythagore.** Si $(x_i)_{i \in I}$ est une famille orthogonale finie d'éléments de E , nous avons $\| \sum_{i \in I} x_i \|^2 = \sum_{i \in I} \|x_i\|^2$.

Définition 29.2

Soient F et G deux sous-espaces vectoriels d'un espace préhilbertien E . Nous disons que F et G sont des **espaces orthogonaux** et nous notons $F \perp G$ lorsque :

$$\forall x \in F, \forall y \in G, \langle x | y \rangle = 0.$$

L'**orthogonal** de F est le sous-espace vectoriel défini par :

$$F^\perp = \{x \in E; \forall y \in F, \langle x | y \rangle = 0\}.$$

Théorème 29.2

Rappelons que E est un espace préhilbertien et F et G deux sous-espaces vectoriels de E . Alors,

$$E^\perp = \{0_E\}; \quad \{0_E\}^\perp = E; \quad F \perp G \iff F \subset G^\perp \iff G \subset F^\perp.$$

$$F \subset G \implies G^\perp \subset F^\perp; \quad F \subset (F^\perp)^\perp; \quad F \cap F^\perp = \{0_E\}.$$

Définition 29.3

Dans un espace préhilbertien E , deux sous-espaces vectoriels F et G sont dits **supplémentaires orthogonaux** dans E , ce qui est noté $E = F \oplus^\perp G$ si $E = F \oplus G$ et $F \perp G$.

Théorème 29.3

Soient F et G deux sous-espaces vectoriels de E . Si F et G sont supplémentaires orthogonaux dans E , nous avons $F^\perp = G$, $G^\perp = F$ d'où $(F^\perp)^\perp = F$.

2. Dimension finie

Nous supposons de surcroît que E est de dimension finie n . E est donc un espace euclidien.

Théorème 29.4

Soit F un sous-espace vectoriel de E . Alors,

$$\dim(F^\perp) = \dim E - \dim F \text{ et } F = (F^\perp)^\perp.$$

Théorème 29.5 Méthode d'orthogonalisation de Schmidt

Soit $\{x_1; \dots; x_n\}$ une famille libre de E . Alors il existe une famille libre orthogonale $\{y_1; \dots; y_n\}$ telle que $\text{Vect}(x_1; \dots; x_n) = \text{Vect}(y_1; \dots; y_n)$. Dans la méthode de Schmidt, elle se construit par récurrence en posant $y_1 = x_1$ puis $y_k = x_k - \sum_{i=1}^{k-1} \lambda_i y_i$, avec $\lambda_i = \frac{\langle y_i | x_k \rangle}{\langle y_i | y_i \rangle}$.

Théorème 29.6

Soit F un sous-espace vectoriel de E . Supposons que $E = F \oplus F^\perp$. Nous pouvons définir le **projecteur orthogonal** p_F sur F . Si $\{e_1; \dots; e_p\}$ est une base orthonormale de F , nous avons :

$$\forall x \in E, p_F(x) = \sum_{i=1}^p \langle e_i | x \rangle e_i.$$

Définition 29.4

Soit F un sous-espace vectoriel de E . $\forall x \in E$, nous appelons **distance** de x au sous-espace de dimension finie F le nombre :

$$d(x, F) = \inf_{z \in F} \|x - z\|.$$

Théorème 29.7

Soit F un sous-espace vectoriel de E . Pour tout $x \in E$, la distance $d(x, F)$ est un minimum atteint en un point et un seul $z = p_F(x)$ tel que $d(x, F) = \|x - p_F(x)\|$ et nous avons l'égalité de Pythagore :

$$\|x\|^2 = \|p_F(x)\|^2 + d(x, F)^2.$$

$p_F(x)$ est appelé le **projecteur orthogonal de x sur F** .

Définition 29.5

Dans un espace euclidien E , un endomorphisme f est dit **orthogonal** s'il vérifie l'une des conditions équivalentes suivantes :

- il conserve le produit scalaire : $\forall x \in E, \forall y \in E, \langle f(x)|f(y) \rangle = \langle x|y \rangle$. Nous disons aussi que f est une **isométrie vectorielle**.
- il conserve la norme : $\forall x \in E, \|f(x)\| = \|x\|$.
- il existe une base orthonormale $\mathcal{B}$ de E telle que $f(\mathcal{B})$ soit une base orthonormale de E .
- pour toute base orthonormale $\mathcal{B}$ de E , $f(\mathcal{B})$ est une base orthonormale de E .

Définition 29.6

Soit E un espace euclidien. Une matrice carrée A d'ordre n est dite **orthogonale** si et seulement si l'une des conditions équivalentes suivantes est vérifiée :

- C'est la matrice de passage d'une base orthonormale $\mathcal{B}$ de E à une base orthonormale $\mathcal{B}'$ de E .
- Ses vecteurs colonnes (ainsi que ses vecteurs lignes) vérifient $\forall i, j \in \{1; \dots; n\}, \langle C_i|C_j \rangle = 1$ si $i = j$, 0 sinon.
- ${}^tAA = I_n$
- A est inversible et ${}^tA = A^{-1}$.

Exemple

Dans $\mathbb{R}^2$, muni de la base $\{\vec{i}; \vec{j}\}$, la matrice de rotation d'angle θ est une matrice orthogonale et s'écrit :

$$\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}.$$

Application

Nous considérons la matrice carrée d'ordre 3 suivante : $A = \begin{pmatrix} x & y & z \\ z & x & y \\ y & x & z \end{pmatrix}$ et posons $a = xy + yz + zx$ et $b = x + y + z$ avec $x, y, z \in \mathbb{R}$.

1. Montrer que A est une matrice orthogonale d'ordre trois si et seulement si $a = 0$ et $b = 1$ ou -1 .
2. Montrer que A appartient au groupe des rotations des matrices orthogonales (c'est-à-dire des matrices orthogonales dont le déterminant vaut 1) d'ordre trois si et seulement si $a = 0$ et $b = 1$.

1.



Cherchons tout d'abord les conditions sur x, y et z qui font de A une matrice orthogonale d'ordre trois. Si A est orthogonale alors les vecteurs colonnes C_1, C_2 et C_3 vérifient, $\forall i = 1, 2, 3, \langle C_i, C_j \rangle = 1$ si $i = j$, 0 sinon.

$$\langle C_1, C_2 \rangle = \langle C_1, C_3 \rangle = \langle C_2, C_3 \rangle = 0 \iff xy + xz + yz = 0 \iff a = 0.$$

D'autre part,

$$\langle C_1, C_1 \rangle = \langle C_2, C_2 \rangle = \langle C_3, C_3 \rangle = 1 \iff x^2 + b^2 + c^2 = 1 \iff (x + y + z)^2 - 2a = 1.$$

Comme $a = 0$, nous en déduisons que $(x + y + z)^2 = 1$ et que donc $b = x + y + z = 1$ ou -1 .

2.



A appartient au groupe des rotations orthogonales d'ordre 3 si et seulement si A est orthogonale et $\det(A) = 1$. Les conditions sur x, y et z pour que A soit orthogonale sont $a = 0$ et $b = 1$ ou $b = -1$. Calculons $\det(A)$.

En sommant les trois lignes de la matrice A et en remplaçant la première par cette somme pour le calcul du déterminant, nous avons :

$$\begin{aligned} \det(A) &= (x + y + z) \times \begin{vmatrix} 1 & 1 & 1 \\ z & x & y \\ y & z & x \end{vmatrix} \\ &= (x + y + z) \times \begin{vmatrix} 1 & 0 & 0 \\ z & x - z & y - z \\ y & z - y & x - y \end{vmatrix} \\ &= (x + y + z) \times [(x - z)(x - y) - (y - z)(z - y)] \\ &= (x + y + z) \times [x^2 + y^2 + z^2 - xy - xz - yz] \\ &= (x + y + z) \times [x^2 + y^2 + z^2 - a] \\ &= (x + y + z) \times [x^2 + y^2 + z^2]^2 \\ &= (x + y + z)^3. \end{aligned}$$

Comme nous devons avoir $\det(A) = 1$, il suffit que $b = x + y + z = 1$.

Pour s'entraîner (solutions p. 518)

29.1 Soit $E = \mathbb{R}^3$. Nous proposons de faire l'étude de l'endomorphisme f de E dont la matrice est :

$$M = \frac{1}{9} \begin{pmatrix} 8 & -1 & -4 \\ -1 & 8 & -4 \\ -4 & -4 & -7 \end{pmatrix}.$$

1. Montrer que M est orthogonale.

2. Montrer que f est une réflexion par rapport au plan (P) d'équation : $x + y + 4z = 0$.

29.2 Soit $\mathbb{R}^3$ muni de la base $\{v_1; v_2; v_3\}$

$$\text{avec } v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, v_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \text{ et } v_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \text{ Utiliser}$$

le procédé d'orthogonalisation de Schmidt pour transformer cette base en une base orthonormale.

29.3 Dans l'espace euclidien $\mathbb{R}^3$, donner les matrices de rotation respectivement autour des axes x, y et z .

29.4 Donner la matrice de rotation qui correspond à une rotation horaire de 90 degrés dans le plan.

Les coniques sont des courbes planes algébriques très utilisées en géométrie.

Définition 30.1

Soit F un point du plan, $\mathcal{D}$ une droite ne passant pas par F , e un réel strictement positif. Pour tout point M du plan, nous notons $d(M, \mathcal{D})$ la distance du point M à la droite $\mathcal{D}$. L'ensemble des points M tels que $\frac{MF}{d(M, \mathcal{D})} = e$ est la **conique** de **foyer** F , de **direction** (ou **directrice**) $\mathcal{D}$ et d'**excentricité** e . C'est une **ellipse** si $e < 1$, une **parabole** si $e = 1$ et une **hyperbole** si $e > 1$. La perpendiculaire à $\mathcal{D}$ passant par F est l'**axe focal** de la conique.

Nous commençons par définir une parabole.

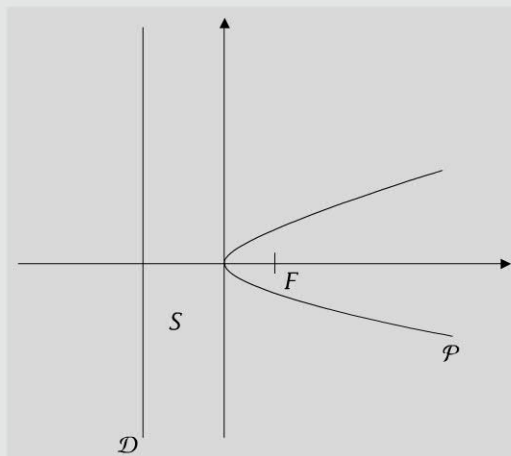
Définition 30.2

La **parabole** $\mathcal{P}$ et son axe focal ont un point commun unique, le **sommet** S . Dans le repère orthonormal d'origine S et admettant l'axe focal comme axe des abscisses, l'équation de la parabole $\mathcal{P}$ est :

$$y^2 = 2px, \text{ pour } x \in \mathbb{R}_+$$

où $p \in \mathbb{R}_+^*$ s'appelle le **paramètre** de la parabole. Le foyer F a pour coordonnées $\left(\frac{p}{2}, 0\right)$. La directrice $\mathcal{D}$ a pour équation $x = -\frac{p}{2}$.

La représentation paramétrique de la parabole est donnée par : $y = 2pt$ et $x = 2pt^2$ ou encore $x = \frac{y^2}{2p}$. Enfin, pour tout point $M_0 \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$ de la parabole, la tangente au point $M_0 \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$ a pour équation $y_0 y = p(x + x_0)$.



Nous poursuivons nos définitions avec celle d'une ellipse.

Définition 30.3

Soit une ellipse ϵ de foyer F , de directrice $\mathcal{D}$ et d'excentricité e . L'**ellipse** ϵ et son axe focal ont deux points communs A et A' , qui sont les **sommets** de l'axe focal. Le milieu O de $[AA']$ est le **centre** de ϵ . La médiatrice de $[AA']$ est l'**axe non focal**. Elle coupe l'ellipse en B et B' , qui sont les **sommets de l'axe non focal**.

Si F' et $\mathcal{D}'$ sont les symétriques de F et $\mathcal{D}$ par rapport à l'origine O du repère orthonormal du plan, l'ellipse de foyer F' , de directrice $\mathcal{D}'$ et d'excentricité e est la même ellipse ϵ . Définissons a et b tels que $AA' = 2a$ et $BB' = 2b$. Dans le repère orthonormal d'origine O et admettant l'axe focal comme axe des abscisses et l'axe non focal comme axe des ordonnées, l'équation de ϵ est :

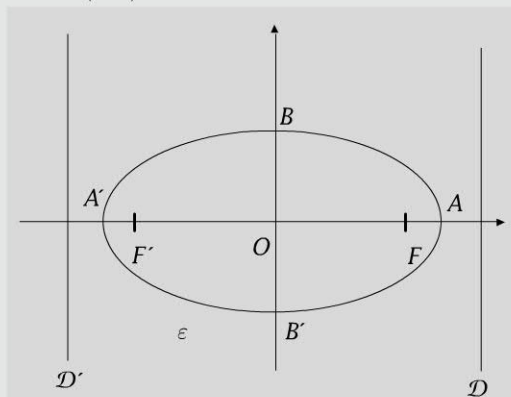
$$\forall x, y \in \mathbb{R}, \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$

En posant $c = \sqrt{a^2 - b^2}$ (possible car nous supposons que $a \geq b$), nous avons $FF' = 2c$, l'excentricité $e = \frac{c}{a}$ et la directrice $\mathcal{D}$ a pour équation $x = \frac{a^2}{c}$.

La représentation paramétrique de l'ellipse est donnée par : $x = a \cos \theta$ et $y = b \sin \theta$ avec $\theta \in [0; 2\pi[$.

Nous pouvons également définir l'ellipse par la **caractérisation bifocale**. Soient deux points distincts du plan F et F' et a un réel positif tel que $FF' < 2a$. L'ensemble des points M du plan tels que $MF + MF' = 2a$ est une ellipse de foyers F et F' .

La tangente à tout point $M_0 \left(\begin{smallmatrix} x_0 \\ y_0 \end{smallmatrix} \right)$ de l'ellipse a pour équation : $\frac{x_0 x}{a^2} + \frac{y_0 y}{b^2} = 1$.



Enfin, nous présentons l'hyperbole.

Définition 30.4

Soit une hyperbole $\mathcal{H}$ de foyer F , de directrice $\mathcal{D}$ et d'excentricité e . L'**hyperbole** $\mathcal{H}$ et son axe focal ont deux points communs, les **sommets** A et A' . Le milieu O de $[AA']$ est le **centre** de $\mathcal{H}$. La médiatrice de $[AA']$ est l'**axe non focal**. Il ne rencontre pas l'hyperbole. Si F' et $\mathcal{D}'$ sont les symétriques de F et $\mathcal{D}$ par rapport à O , l'hyperbole de foyer F' , de directrice $\mathcal{D}'$ et d'excentricité e est la même hyperbole $\mathcal{H}$. Soient $a, b, c \in \mathbb{R}^+$ tels que $AA' = 2a$, $FF' = 2c$ et $b^2 = c^2 - a^2$. Dans le repère orthonormal

d'origine O et admettant l'axe focal comme axe des abscisses et l'axe non focal comme axe des ordonnées, l'équation de $\mathcal{H}$ est :

$$\forall x, y \in \mathbb{R}, \frac{x^2}{a^2} - \frac{y^2}{b^2} = 1.$$

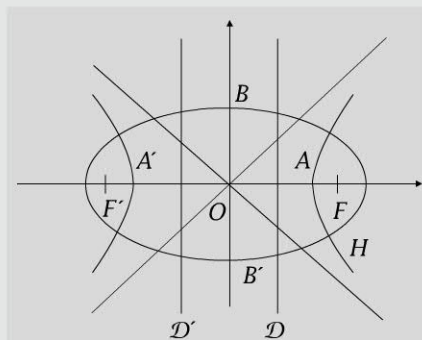
Nous avons l'excentricité $e = \frac{c}{a}$ et la directrice $\mathcal{D}$ a pour équation $x = \frac{a^2}{c}$.

Les **asymptotes** ont pour équation $y = \frac{b}{a}x$ et $y = -\frac{b}{a}x$.

La représentation paramétrique de l'hyperbole est donnée par $x = a \operatorname{ch}(\theta)$ et $y = b \operatorname{sh}(\theta)$ avec $\theta \in \mathbb{R}$.

Nous pouvons également définir l'hyperbole par sa caractérisation bifocale. Soient F et F' deux points distincts du plan et a un réel positif tel que $0 < 2a < FF'$. L'ensemble des points M du plan tels que $|MF - MF'| = 2a$ est une hyperbole de foyers F et F' .

Enfin, la tangente au point $M_0 \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$ a pour équation : $\frac{x_0 x}{a^2} - \frac{y_0 y}{b^2} = 1$.



Application

Nous voulons donner la nature et les caractéristiques de la conique dont l'équation dans le repère orthonormé $(O; \vec{i}; \vec{j})$ est donnée par :

$$y^2 - 6y + 2x + 10 = 0.$$



Tout d'abord, essayons de reconnaître le type de conique dont il s'agit.

Nous avons :

$$\begin{aligned} y^2 - 6y + 2x + 10 = 0 &\iff y^2 - 6y + 10 = -2x, \\ &\iff (y - 3)^2 - 9 + 10 = -2x, \\ &\iff (y - 3)^2 + 1 = -2x, \\ &\iff (y - 3)^2 = -2x - 1, \\ &\iff (y - 3)^2 = -2\left(x + \frac{1}{2}\right). \end{aligned}$$

Nous reconnaissons bien l'équation d'une parabole de la forme $Y^2 = -2X$.

La directrice D a pour équation $x = \frac{1}{2}$ et le foyer F a pour coordonnées $F \begin{pmatrix} -\frac{1}{2} \\ 0 \end{pmatrix}$.

Pour s'entraîner (solutions p. 519)

30.1 Déterminer une équation cartésienne de la parabole de foyer $F\left(\frac{1}{2}\right)$ et de directrice

$$\mathcal{D} : x = 3.$$

30.2 Donner la nature et les éléments de la courbe d'équation dans le repère orthonormé $(O; \vec{i}; \vec{j})$:

$$16x^2 - 24xy + 9y^2 + 35x - 20y = 0.$$

30.3 Dans le repère orthonormé $(O; \vec{i}; \vec{j})$, nous considérons la courbe d'équation suivante :

$$-3x^2 + y^2 - 8x - 4 = 0.$$

1. Montrer qu'il s'agit d'une hyperbole.

2. Déterminer le foyer, la directrice et l'excentricité de cette conique.

30.4 Donner la nature et les éléments de la courbe d'équation dans le repère orthonormé $(O; \vec{i}; \vec{j})$:

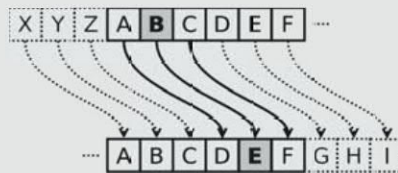
$$2x^2 + 3y^2 - 4x + 18y + 11 = 0.$$

30.5 Soient $\alpha \in \mathbb{R}$ et C_λ l'ensemble des points $M(x; y)$ tels que :

$$x^2 + y^2 + 2\lambda xy - 1 = 0.$$

1. Discuter de la nature de la conique en fonction de λ .

2. Préciser C_0 , C_1 et C_{-1} .



La cryptographie, branche de la cryptologie, est une discipline permettant, au travers de nombreuses méthodes, de protéger des messages à caractère confidentiel. Elle s'utilise dans de nombreux domaines tels que les domaines bancaire, industriel ou encore militaire pour n'en citer que quelques uns.

Nous allons, au travers d'un exemple, faire l'étude d'une de ces méthodes en utilisant les propriétés de l'arithmétique et des matrices. Il s'agit de transformer un message, à l'aide d'une clé de chiffrement, en un message codé qui reste incompréhensible si nous ne connaissons pas la clé de déchiffrement. Le chiffre de Hill, étudié par Lester S. Hill, permet de chiffrer le message en substituant les lettres de celui-ci par des groupes de lettres, ce qui permet de rendre plus difficile le cassage du code. Nous proposons de coder le message suivant :

ALGÈBRE LINÉAIRE

Chaque caractère est d'abord codé par un nombre compris entre 0 et 25. Les lettres de l'alphabet sont associées à ces nombres de la manière suivante :

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20

V	W	X	Y	Z
21	22	23	24	25

Le message ALGEBRE LINEAIRE devient alors la suite de chiffres suivante :

0; 11; 6; 4; 1; 17; 4; 11; 8; 13; 4; 0; 8; 17; 4

Ces chiffres, que nous pouvons considérer comme des éléments de $\mathbb{Z}/26\mathbb{Z}$, vont alors être regroupés par blocs de deux pour former ici les 8 vecteurs suivants (les chiffres étant en nombre impair, nous ajoutons un 0 dans le dernier vecteur) :

$$x_1 = \begin{pmatrix} 0 \\ 11 \end{pmatrix}; x_2 = \begin{pmatrix} 6 \\ 4 \end{pmatrix}; x_3 = \begin{pmatrix} 1 \\ 17 \end{pmatrix}; x_4 = \begin{pmatrix} 4 \\ 11 \end{pmatrix};$$

$$x_5 = \begin{pmatrix} 8 \\ 13 \end{pmatrix}; x_6 = \begin{pmatrix} 4 \\ 0 \end{pmatrix}; x_7 = \begin{pmatrix} 8 \\ 17 \end{pmatrix}; x_8 = \begin{pmatrix} 4 \\ 0 \end{pmatrix}.$$

Nous choisissons ensuite une matrice A de taille 2×2 qui sera la clé de chiffrement. Nous prenons par exemple, la matrice A proposée dans l'article « Matrices et codes secret » de la revue *Tangente*, hors série (num 44, août 2012) intitulée « Les matrices.

Une représentation du monde », $A = \begin{pmatrix} 2 & 5 \\ 1 & 20 \end{pmatrix}$.

Nous avons $\det(A) = 35$ et $A^{-1} = \frac{1}{35} \begin{pmatrix} 20 & -5 \\ -1 & 2 \end{pmatrix}$.

Comme nous travaillons dans $\mathbb{Z}/26\mathbb{Z}$ qui n'est pas un corps, le lien entre les notions de déterminant et d'inversibilité de matrices que nous avons rencontré par exemple dans la Fiche 25 s'en trouve quelque peu modifié. En effet, dans ce cas, une matrice est inversible si et seulement si son déterminant est inversible en tant qu'élément de $\mathbb{Z}/26\mathbb{Z}$. Comme $3 \times 35 = 105$ et comme le reste de la division de 105 par 26 est 1, alors l'inverse de 35 dans $\mathbb{Z}/26\mathbb{Z}$ vaut 3 et donc dans $\mathbb{Z}/26\mathbb{Z}$, $A^{-1} = 3 \begin{pmatrix} 20 & -5 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 8 & 11 \\ 23 & 6 \end{pmatrix}$.

Il ne reste plus qu'à effectuer le produit de la matrice A par chacun des vecteurs précédemment construits pour obtenir, dans $\mathbb{Z}/26\mathbb{Z}$, les vecteurs suivants :

$$A \times x_1 = \begin{pmatrix} 55 \\ 220 \end{pmatrix} = \begin{pmatrix} 3 \\ 12 \end{pmatrix} = y_1; \quad A \times x_2 = \begin{pmatrix} 32 \\ 86 \end{pmatrix} = \begin{pmatrix} 6 \\ 8 \end{pmatrix} = y_2$$

$$A \times x_3 = \begin{pmatrix} 87 \\ 341 \end{pmatrix} = \begin{pmatrix} 9 \\ 3 \end{pmatrix} = y_3; \quad A \times x_4 = \begin{pmatrix} 63 \\ 224 \end{pmatrix} = \begin{pmatrix} 11 \\ 16 \end{pmatrix} = y_4$$

$$A \times x_5 = \begin{pmatrix} 81 \\ 268 \end{pmatrix} = \begin{pmatrix} 3 \\ 8 \end{pmatrix} = y_5; \quad A \times x_6 = \begin{pmatrix} 8 \\ 4 \end{pmatrix} = y_6$$

$$A \times x_7 = \begin{pmatrix} 101 \\ 348 \end{pmatrix} = \begin{pmatrix} 23 \\ 10 \end{pmatrix} = y_7; \quad A \times x_8 = \begin{pmatrix} 8 \\ 4 \end{pmatrix} = y_8.$$

À partir de ces 8 nouveaux vecteurs, nous obtenons la suite de chiffres suivante :

3; 12; 6; 8; 9; 3; 11; 16; 3; 8; 8; 4; 23; 10; 8; 4

Si nous codons ces 16 chiffres par des lettres, nous obtenons le texte suivant :

DMGIJDLQDIIEKIE

Pour déchiffrer ce dernier message, utilisons le raisonnement inverse, c'est-à-dire multiplions la matrice A^{-1} , dite de déchiffrement, par chacun des vecteurs $y_1, y_2, \dots, y_8$, puis utilisons le codage des chiffres ainsi obtenus par les lettres de l'alphabet.

Analyse

Partie

2



Introduction

L'analyse est le langage mathématique qui permet de résoudre de nombreux problèmes dans les différents domaines des sciences de l'ingénieur. Nous traitons ici les outils essentiels avec un minimum d'abstraction. Nous commençons par les nombres réels et les suites (fiches 31 à 35) qui permettent d'étudier des phénomènes de la vie quotidienne. Puis nous enchaînons avec l'étude d'une fonction d'une variable (fiches 36 à 57).

Nous poursuivons avec les fonctions de plusieurs variables (fiches 58 à 68) souvent utilisées pour modéliser des phénomènes réels. Enfin nous terminerons avec la transformée de Laplace (fiche 69) et une introduction aux équations aux dérivées partielles (fiche 70), notions incontournables pour des élèves ingénieurs.

Un focus sur le wifi illustre cette partie.

Les bonus web sur Dunod.com

 Retrouvez les corrigés détaillés d'une sélection d'exercices sur le site dunod.com

1. Définitions fondamentales

Définition 31.1

L'ensemble $\mathbb{R}$, appelé **ensemble des nombres réels**, est muni de deux lois internes $+$, $\times$, et d'une relation $\leq$, tel que :

- $(\mathbb{R}, +, \times)$ est un corps commutatif ;
- $\leq$ est une relation d'ordre total dans $\mathbb{R}$;
- $\forall (x, y, z) \in \mathbb{R}^3, x \leq y \implies x + z \leq y + z$ et $\forall (x, y) \in \mathbb{R}^2, \forall z \in [0; +\infty[, x \leq y \implies xz \leq yz$.

 Fiche 8

Définition 31.2

- Une partie P de $\mathbb{R}$ est dite **majorée** s'il existe un réel M tel que pour tout $x \in P, x \leq M$. M est alors un **majorant** de P dans $\mathbb{R}$.
- Une partie P de $\mathbb{R}$ est dite **minorée** s'il existe un réel m tel que pour tout $x \in P, m \leq x$. m est alors un **minorant** de P dans $\mathbb{R}$.
- Une partie P de $\mathbb{R}$ est dite **bornée** si et seulement si P est majorée et minorée.

 Fiche 4

Définition 31.3

- La **borne supérieure** de P dans $\mathbb{R}$, notée $\sup(P)$, est le plus petit élément (s'il existe) de l'ensemble des majorants de P dans $\mathbb{R}$.
- La **borne inférieure** de P dans $\mathbb{R}$, notée $\inf(P)$, est le plus grand élément (s'il existe) de l'ensemble des minorants de P dans $\mathbb{R}$.

 Fiche 4



Si la partie P admet un plus grand élément, alors c'est la borne supérieure de P . Si la partie P admet un plus petit élément, alors c'est la borne inférieure de P .

Théorème 31.1

Toute partie non vide et majorée (resp. minorée) de $\mathbb{R}$ admet une borne supérieure (resp. inférieure).

2. Valeur absolue d'un réel

Définition 31.4

La **valeur absolue d'un réel** x , notée $|x|$, est définie par : $|x| = x$ si $x \geq 0$, $|x| = -x$ si $x \leq 0$.

Proposition 31.1

Nous avons les propriétés suivantes :

- $\forall x \in \mathbb{R}, |x| = \max(x, -x)$.
- $\forall x \in \mathbb{R}, |x| \geq 0$.
- $\forall x \in \mathbb{R}, |x| = 0 \iff x = 0$.
- $\forall (x, y) \in \mathbb{R}^2, |xy| = |x||y|$.
- $\forall (x, y) \in \mathbb{R}^2, |x| - |y| \leq |x + y| \leq |x| + |y|$, **inégalité triangulaire**.
- $\forall (x, y) \in \mathbb{R}^2, ||x| - |y|| \leq |x - y|$, **inégalité triangulaire renversée**.

Définition 31.5

Soient deux réels x et y . Nous appelons **distance** de x à y , et nous notons $d(x, y)$, le nombre réel positif défini par $d(x, y) = |x - y|$.

Proposition 31.2

La distance vérifie les propriétés suivantes :

- Positivité : $\forall (x, y) \in \mathbb{R}^2, d(x, y) \geq 0$.
- Séparation : $\forall (x, y) \in \mathbb{R}^2, d(x, y) = 0 \iff x = y$.
- Symétrie : $\forall (x, y) \in \mathbb{R}^2, d(x, y) = d(y, x)$.
- Inégalité triangulaire : $\forall (x, y, z) \in \mathbb{R}^3, d(x, z) \leq d(x, y) + d(y, z)$.

3. Intervalles**Définition 31.6**

Soient a, b deux réels tels que $a < b$.

- Soient I un intervalle non vide et borné, a la borne inférieure de I et b la borne supérieure de I . Alors I est l'un des intervalles : $[a; b]$, $]a; b[$, $[a; b[$ et $]a; b]$. L'intervalle $[a; b]$ est dit **intervalle fermé borné** ou encore **segment**. L'intervalle $]a; b[$ est dit intervalle **ouvert**. Les intervalles $[a; b[$, $]a; b]$ sont appelés intervalles **semi-ouverts**, ou intervalles **semi-fermés**.
- Soient I un intervalle non vide, majoré mais non minoré, b la borne supérieure de I . Alors I est l'un des intervalles : $] - \infty; b]$ (intervalle fermé) ou $] - \infty; b[$ (intervalle ouvert).
- Soient I un intervalle non vide, minoré mais non majoré, a la borne inférieure de I . Alors I est l'un des intervalles : $[a; +\infty[$ (intervalle fermé) ou $]a; +\infty[$ (intervalle ouvert).
- Soit I un intervalle non vide, ni minoré et ni majoré. Alors $I =] - \infty; +\infty[= \mathbb{R}$.

4. Approximation d'un réel

Définition 31.7

Une partie P de $\mathbb{R}$ est dite **dense dans** $\mathbb{R}$ si et seulement si :

$$\forall (x, y) \in \mathbb{R}^2, \quad x < y \implies (\exists p \in P \text{ tel que } x < p < y).$$

ou encore si et seulement si :

$$\forall x \in \mathbb{R}, \forall \epsilon > 0, \exists p \in P, \text{ tel que } |x - p| < \epsilon.$$

Théorème 31.2

$\mathbb{Q}$ et son complémentaire $\mathbb{R} \setminus \mathbb{Q}$ sont **denses** dans $\mathbb{R}$.

Définition 31.8

Étant donné un nombre réel x , il existe un unique entier relatif, noté $E(x)$ ou $[x]$, tel que $E(x) \leq x < E(x) + 1$. $E(x)$ est appelé la **partie entière** de x .

Théorème 31.3

Soient $x \in \mathbb{R}$ et $p \in \mathbb{N}$. Nous avons :

$$E(x \times 10^p) \times 10^{-p} \leq x < E(x \times 10^p) \times 10^{-p} + 10^{-p}.$$

$E(x \times 10^p) \times 10^{-p}$ est la **valeur décimale approchée de x à 10^{-p} près par défaut** et $(E(x \times 10^p) + 1) \times 10^{-p}$ est la **valeur décimale approchée de x à 10^{-p} près par excès**.

Application **Bornes supérieure et inférieure**

1. Calculer, si elles existent, les bornes supérieure et inférieure de la partie P de $\mathbb{R}$ définie par : $P = \left\{ \sin\left(\frac{2n\pi}{7}\right), n \in \mathbb{Z} \right\}$.

2. Soit F la partie de $\mathbb{R}$ définie par : $F = \left\{ \frac{1}{p^2} + \frac{1}{q^2} - \frac{1}{pq}, (p, q) \in (\mathbb{N}^*)^2 \right\}$. Montrer que F admet une borne inférieure m et une borne supérieure M et les calculer.

1. Nous avons pour tout $x \in \mathbb{R}$, $\sin(x) \in [-1; 1]$ donc $P \subset [-1; 1]$. La partie P est non vide et bornée. Donc la partie P admet une borne supérieure, notée $\sup(P)$ et une borne inférieure, notée $\inf(P)$. De plus, $\sup P \leq 1$ et $\inf P \geq -1$. La fonction $f : n \mapsto \sin\left(\frac{2n\pi}{7}\right)$ est impaire et périodique de période 7. Par conséquent $\sup(P) = \sin\left(\frac{4\pi}{7}\right)$ et $\inf(P) = \sin\left(\frac{10\pi}{7}\right) = -\sin\left(\frac{4\pi}{7}\right)$.

2. D'une part, pour tout $(p, q) \in (\mathbb{N}^*)^2$, nous avons : $\frac{1}{p^2} + \frac{1}{q^2} - \frac{1}{pq} = \left(\frac{1}{p} - \frac{1}{q}\right)^2 + \frac{1}{pq} > 0$.
Donc F est minorée par 0 et possède une borne inférieure $m \geq 0$.

D'autre part, nous avons, comme $p \geq 1$ et $q \geq 1$, $1 - \frac{1}{p^2} \geq 0$ et $1 - \frac{1}{q^2} \geq 0$. Ainsi $\left(1 - \frac{1}{p^2}\right) \times \left(1 - \frac{1}{q^2}\right) \geq 0$ et en développant nous avons : $\frac{1}{p^2} + \frac{1}{q^2} \leq 1 + \frac{1}{p^2 q^2}$. De plus $\frac{1}{p^2 q^2} \leq \frac{1}{pq}$, donc $\frac{1}{p^2} + \frac{1}{q^2} - \frac{1}{pq} \leq 1$. Donc F est majorée par 1 et possède une borne supérieure $M \leq 1$.

Déterminons les deux bornes maintenant. Pour $p = q \in \mathbb{N}^*$, $\frac{1}{p^2} + \frac{1}{q^2} - \frac{1}{pq} = \frac{1}{p^2} \rightarrow 0$ quand $p \rightarrow +\infty$. Ainsi F possède un minorant qui est nécessairement inférieur ou égal à 0. La borne inférieure de F est donc 0.

Si $p = q = 1$, alors $\frac{1}{p^2} + \frac{1}{q^2} - \frac{1}{pq} = 1$ donc 1 est un élément de F . Or nous avons vu que 1 majore F . C'est donc la borne supérieure de F .

Pour s'entraîner (solutions p. 519)

31.1 Une équation à résoudre

Résoudre dans $\mathbb{R}$ l'équation :

$$E(\sqrt{x}) = \sqrt{E(x)}.$$

31.2 Une égalité remarquable

Montrer que pour tout $n \in \mathbb{N}^*$ et pour tout

$$x \in \mathbb{R} : E\left(\frac{1}{n}E(nx)\right) = E(x).$$

31.3 Montrer que pour tout $x \in \mathbb{R}$:

$$E\left(\frac{x}{2}\right) + E\left(\frac{x+1}{2}\right) = E(x).$$

31.4 Soient F_1 , F_2 et F_3 les parties de $\mathbb{R}$ définies par :

$$F_1 = \{x \ln x + y \ln y, \ x + y = 1, \ x, y \in \mathbb{R}_+^*\};$$

$$F_2 = \left\{\frac{pq}{p^2 + q^2 + 1}, \ p, q \in \mathbb{R}^*\right\};$$

$$F_3 = \{x^2 + y^2, \ xy = 1, \ x, y \in \mathbb{R}\}.$$

Démontrer que F_1 et F_2 possèdent une borne supérieure et une borne inférieure puis les calculer. Démontrer que F_3 possède une borne inférieure et la calculer. F_3 possède-t-elle une borne supérieure ?

1. Suite numérique

Définition 32.1

- Une **suite numérique** est une application de $\mathbb{N}$ dans $\mathbb{K}$, où $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$; au lieu de la noter

$$\begin{aligned} u : \mathbb{N} &\longrightarrow \mathbb{K} \\ n &\longmapsto u(n) \end{aligned}$$

nous la notons souvent $(u_n)_{n \in \mathbb{N}}$ ou $(u_n)_{n \geq 0}$ ou même encore $(u_n)_n$.

- Une **suite réelle** (respectivement **complexe**) est une suite numérique telle que : $\forall n \in \mathbb{N}, u_n \in \mathbb{R}$ (respectivement $\mathbb{C}$).



- Les suites sont en fait souvent définies sur une partie infinie I de $\mathbb{N}$, par exemple sur $\mathbb{N}^*$. Il convient alors d'adapter les définitions précédentes et suivantes à cette situation.
- **Attention :** La notation (u_n) désigne la suite en tant qu'objet mathématique, parfois notée simplement u ou $(u_n)_{n \in I}$ lorsqu'il est nécessaire de préciser la partie I de $\mathbb{N}$ sur laquelle la suite est définie, tandis que la notation u_n désigne le terme général de la suite (u_n) .

Une suite peut être définie de plusieurs manières, nous allons exposer les plus fréquentes ci-dessous.

2. Comment définir une suite ?

► Suite définie explicitement

Dans ce cas, nous connaissons explicitement l'application u de $\mathbb{N}$ dans $\mathbb{R}$.

Exemple

Soit la suite de terme général $u_n = \frac{1}{n}$ définie pour $n > 0$. Les termes successifs sont : $1, \frac{1}{2}, \dots, \frac{1}{n}, \dots$. Cette suite est définie explicitement en fonction de n .

► Suite définie par une relation de récurrence

Définition 32.2

Une **suite récurrente** est définie par la donnée d'un ou plusieurs termes de la suite (souvent le premier terme) et d'une relation liant deux (ou plusieurs) termes consécutifs u_n et u_{n+1} . La suite se définit donc par la relation de récurrence : $u_{n+1} = \varphi(u_n)$ ($u_{n+p} = \varphi(u_n, u_{n+1}, \dots, u_{n+p-1})$), où φ est une fonction explicite connue, et la condition initiale $u_0 = a$, où $a \in \mathbb{R}$ ($u_0, \dots, u_{p-1} \in \mathbb{R}$).

► Suite définie implicitement par une propriété

Exemple

Pour tout $n \in \mathbb{N}$, u_n est l'unique solution de l'équation $u_n \exp(u_n) = n$.

Dans toute la suite, nous considérons des suites de nombres réels.

3. Suites réelles monotones

Définition 32.3

- La suite (u_n) est une **suite croissante (strictement croissante)** si :

$$\forall n \in \mathbb{N}, \quad u_{n+1} \geq (>) u_n.$$

- La suite (u_n) est une **suite décroissante (strictement décroissante)** si :

$$\forall n \in \mathbb{N}, \quad u_{n+1} \leq (<) u_n.$$

- La **suite** (u_n) est **monotone** si la suite (u_n) est croissante ou décroissante.
- La **suite** (u_n) est **strictement monotone** si la suite (u_n) est strictement croissante ou strictement décroissante.

4. Suites bornées

Définition 32.4

- Une **suite réelle** (u_n) est **majorée** s'il existe un réel M tel que pour tout entier naturel n , $u_n \leq M$. Le réel M est appelé un **majorant** de la suite réelle. Dès lors qu'une suite est majorée, il existe une infinité de majorants (tous les réels supérieurs à un majorant quelconque).
- Une **suite réelle** (u_n) est **minorée** s'il existe un réel m tel que pour tout entier naturel n , $u_n \geq m$. Le réel m est appelé un **minorant** de la suite réelle. Dès lors qu'une suite est minorée, il existe une infinité de minorants (tous les réels inférieurs à un minorant quelconque).
- Une **suite réelle** (u_n) est **bornée** si (u_n) est à la fois **majorée** et **minorée**.

5. Suites particulières

Définition 32.5

Une **suite** (u_n) est **alternée** si deux termes consécutifs quelconques sont de signes opposés : $u_n u_{n+1} \leq 0$, pour tout $n \in \mathbb{N}$.

Définition 32.6

Une **suite** (u_n) est **périodique** de période $p \in \mathbb{N}^*$ si pour tout $n \in \mathbb{N}$, $u_{n+p} = u_n$.

Exemple

Soit la suite de terme général $u_n = (-1)^n$, définie pour tout $n \in \mathbb{N}$. Cette suite est périodique de période 2. En effet, nous avons $u_{n+2} = (-1)^{n+2} = (-1)^n = u_n$. De plus, cette suite est aussi alternée car $u_n = 1$ lorsque n est pair et $u_n = -1$ lorsque n est impair.

Définition 32.7

Une **suite** (u_n) est **stationnaire** si à partir d'un certain rang, la suite (u_n) est constante.

Application

1. Démontrer que la suite $(u_n)_{n \in \mathbb{N}^*}$ de terme général $u_n = 7 + \frac{3}{n}$ est strictement décroissante et minorée par 7.

2. Démontrer que la suite $(u_n)_{n \in \mathbb{N}^*}$ de terme général $u_n = \frac{1 \cdot 1 + 2 \cdot 2! + \cdots + n \cdot n!}{(n+1)!}$ est strictement croissante.

3. Nous représentons un atome d'une manière approchée par un noyau central sphérique de rayon a avec $a \geq 1$ et de centre O , contenant Z protons, chacun possédant une charge élémentaire $+e > 0$. La charge q_n du cortège électronique de l'atome est donnée par, A étant une constante réelle non nulle :

$$q_n = 4\pi A \frac{a^{3-n}}{n-3} \quad \text{où } n > 3.$$

Le champ électrique E_n à une distance $r > a$ du centre O de l'atome est donné par :

$$E_n = \frac{Zer^{1-n}}{4\pi\epsilon_0 a^{3-n}} \quad \text{où } \epsilon_0 \text{ est la permittivité du vide.}$$

Étudier la monotonie des suites $(q_n)_{n \in \mathbb{N} \setminus \{0,1,2,3\}}$ et $(E_n)_{n \in \mathbb{N}}$.

1. Pour déterminer le sens de variation de la suite, nous calculons la différence entre u_{n+1} et u_n .

$$u_{n+1} - u_n = 7 + \frac{3}{n+1} - 7 - \frac{3}{n} = \frac{-3}{n(n+1)}.$$

Or nous avons $n(n+1) > 0$ et par conséquent $\frac{-3}{n(n+1)} < 0$.

Ainsi $u_{n+1} - u_n < 0$ et nous concluons que la suite $(u_n)_{n \in \mathbb{N}^*}$ est strictement décroissante pour tout $n \in \mathbb{N}^*$.

De plus pour tout $n > 0$, $\frac{3}{n} > 0$. Donc pour tout $n > 0$, $u_n > 7$. Nous en concluons que la suite $(u_n)_{n \in \mathbb{N}^*}$ est minorée par 7.

2.



Nous allons d'abord étudier le numérateur de l'expression du terme général u_n . Notons $v_n = 1 \cdot 1 + 2 \cdot 2! + \cdots + n \cdot n!$ ce numérateur.

Nous avons :

$$v_1 = 1 \cdot 1! = 2! - 1, \quad v_2 = 5 = 3! - 1, \quad v_3 = 23 = 4! - 1.$$

Soit, pour tout $n \geq 1$, (P_n) la propriété $v_n = (n+1)! - 1$. Nous allons démontrer que cette propriété est vraie pour tout $n \geq 1$ par récurrence sur n . Nous avons déjà vérifié que cette propriété est vraie pour $n = 1, 2, 3$. Supposons que la propriété (P_k) est vraie pour un $k \geq 1$ et montrons que (P_k) reste vraie pour $k+1$. Nous avons :

$$v_{k+1} = 1 \cdot 1 + 2 \cdot 2! + \dots + k \cdot k! + (k+1) \cdot (k+1)! = (k+1)! - 1 + (k+1) \cdot (k+1)! = (k+2)! - 1.$$

Donc la propriété (P_{k+1}) est vraie. Ainsi (P_n) est vraie pour tout $n \geq 1$. Nous pouvons alors écrire : $u_n = \frac{(n+1)! - 1}{(n+1)!} = 1 - \frac{1}{(n+1)!}$. La différence $u_{n+1} - u_n = \frac{n+1}{(n+2)!}$ est strictement positive pour tout $n \geq 1$ et par conséquent la suite $(u_n)_{n \in \mathbb{N}^*}$ est strictement croissante.

3. Nous avons, pour $n > 3$:

$$q_{n+1} = 4\pi A \frac{a^{2-n}}{n-2} > 0.$$

Par conséquent, nous pouvons calculer le rapport de deux termes consécutifs qui vaut :

$$\frac{q_{n+1}}{q_n} = \frac{n-3}{a(n-2)}.$$

Or $a \geq 1$, donc $\frac{1}{a} \leq 1$, et $\frac{n-3}{n-2} < 1$ donc $\frac{q_{n+1}}{q_n} < 1$ et ainsi la suite $(q_n)_{n \in \mathbb{N} \setminus \{0,1,2,3\}}$ est strictement décroissante.

D'autre part, nous avons :

$$E_{n+1} = \frac{Zer^{-n}}{4\pi\epsilon_0 a^{2-n}} > 0.$$

Par conséquent, nous pouvons calculer le rapport de deux termes consécutifs qui vaut :

$$\frac{E_{n+1}}{E_n} = \frac{a}{r} < 1.$$

Donc la suite $(E_n)_{n \in \mathbb{N}}$ est strictement décroissante.

Pour s'entraîner (solutions p. 520)

32.1 Soit la suite (u_n) définie par : $u_1 = -\frac{1}{2}$ et $\forall n > 1 \quad u_{n+1} = u_n^2 + \frac{1}{4}$. Montrer par récurrence, pour tout $n > 1$, que : $u_n = \frac{1}{2}$ et en déduire que la suite est stationnaire.

32.2 Étudier la monotonie de la suite (u_n) définie par : $\forall n \in \mathbb{N} \quad u_n = \frac{e^n - 1}{e^n + 1}$.

32.3 Étudier la monotonie de la suite (u_n) définie par : $\forall n \in \mathbb{N} \quad u_n = -2 \left(\frac{2}{3}\right)^n$.

Convergence et divergence d'une suite numérique

1. Notion de convergence

Définition 33.1

- La suite numérique (u_n) converge vers $l \in \mathbb{K}$, où $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, si et seulement si :

$$\forall \epsilon > 0, \quad \exists n_0 \in \mathbb{N} \quad \text{tel que} \quad \forall n \in \mathbb{N}, (n \geq n_0 \implies |u_n - l| \leq \epsilon).$$

Nous écrivons alors $\lim_{n \rightarrow +\infty} u_n = l$.

- Une suite numérique (u_n) diverge si et seulement si elle ne converge pas.
- Une suite réelle diverge vers $+\infty$ si : $\forall M > 0, \exists n_0 \in \mathbb{N}$ tel que $\forall n \geq n_0, u_n \geq M$.
- Une suite réelle diverge vers $-\infty$ si : $\forall m > 0, \exists n_0 \in \mathbb{N}$ tel que $\forall n \geq n_0, u_n \leq -m$.



Il est très important de comprendre le principe de cette définition qui est utilisé également dans le cas des fonctions. L'idée est qu'il est possible de rendre l'écart entre les termes de la suite (u_n) et le nombre l aussi petit que nous le souhaitons, inférieur à tout nombre réel $\epsilon > 0$, à condition de ne considérer que les termes de la suite (u_n) à partir d'un certain rang n_0 . Ainsi, à chaque valeur de ϵ , correspond un rang n_0 à partir duquel tous les termes de la suite sont à une distance de l inférieure à ϵ .

Proposition 33.1

- Si une suite numérique (u_n) converge, alors sa limite l est unique.
- Si une suite réelle (u_n) converge, alors la suite (u_n) est bornée.



La réciproque de la dernière proposition est fausse, comme le montre l'exemple de la suite réelle de terme général $u_n = (-1)^n$. Celle-ci est minorée par -1 et majorée par $+1$, donc bornée mais ne converge pas !

2. Calcul de limites

Théorème 33.1 Limites de référence

Pour des réels $k > 1, \alpha > 0, \beta > 0$,

$$\lim_{n \rightarrow +\infty} \frac{k^n}{n!} = 0; \quad \lim_{n \rightarrow +\infty} \frac{n^\alpha}{k^n} = 0; \quad \lim_{n \rightarrow +\infty} \frac{(\ln n)^\beta}{n^\alpha} = 0.$$

Théorème 33.2 Opérations algébriques des suites convergentes

Soient deux suites (u_n) et (v_n) qui convergent respectivement vers l_1 et vers l_2 . Alors les suites $(u_n + v_n)$, (λu_n) , avec $\lambda \in \mathbb{R}$ et $(u_n v_n)$ convergent respectivement vers $l_1 + l_2$, λl_1 et $l_1 l_2$. Si $l_2 \neq 0$, la suite $\left(\frac{u_n}{v_n}\right)$ converge vers $\frac{l_1}{l_2}$. Si la suite (u_n) converge vers 0 et si la suite (v_n) est bornée, alors la suite $(u_n v_n)$ converge vers 0.

Proposition 33.2 Relation d'ordre

Si (u_n) et (v_n) sont des suites convergentes telles que pour $n \geq n_0, u_n \leq v_n$, alors nous avons $\lim_{n \rightarrow +\infty} u_n \leq \lim_{n \rightarrow +\infty} v_n$.

3. Théorèmes de convergence

Théorème 33.3 Théorème d'encadrement

Si, à partir d'un certain rang $n_0 \in \mathbb{N}$, la relation suivante est vérifiée pour tout $n \geq n_0$: $u_n \leq v_n \leq w_n$ et si les suites (u_n) et (w_n) convergent vers la même limite l , alors la suite (v_n) converge aussi vers l .



Ce théorème est également appelé théorème « des gendarmes ».

Définition 33.2

La suite (v_n) est **extraite** de la suite (u_n) s'il existe une application ϕ de $\mathbb{N}$ dans $\mathbb{N}$, strictement croissante, telle que $v_n = u_{\phi(n)}$. Nous disons aussi que (v_n) est une **sous-suite** de (u_n) .

Théorème 33.4

Si (u_n) converge vers l , alors toute suite extraite de (u_n) converge aussi vers l .

Théorème 33.5 Théorème de Bolzano-Weierstrass

De toute suite de réels bornée, nous pouvons extraire une sous-suite qui converge.

Théorème 33.6 Convergence des suites monotones

- Toute suite de réels majorée et croissante est convergente.
- Toute suite de réels minorée et décroissante est convergente.
- Si une suite est croissante et non majorée, elle diverge vers $+\infty$.
- Si une suite est décroissante et non minorée, elle diverge vers $-\infty$.

Définition 33.3

Deux suites réelles (u_n) et (v_n) sont **adjacentes** si l'une des suites est croissante, l'autre suite décroissante et la suite $(v_n - u_n)$ converge vers 0.

Théorème 33.7

Si deux suites réelles sont adjacentes, alors elles convergent vers la même limite l .

4. Suites complexes

Définition 33.4 Rappel

Une **suite de nombres complexes** est une application u de $\mathbb{N}$ dans $\mathbb{C}$:

$$\begin{aligned} u : \mathbb{N} &\longrightarrow \mathbb{C} \\ n &\longmapsto u(n) = z_n = x_n + iy_n, \\ &\text{où } x_n \text{ et } y_n \text{ sont des termes généraux de deux suites de réels.} \end{aligned}$$

Définition 33.5

La suite (z_n) est **bornée** s'il existe $M \in \mathbb{R}^+$ tel que pour tout $n \in \mathbb{N}$, $|z_n| \leq M$, où $|z|$ désigne ici le module d'un nombre complexe z .



Pour une suite complexe, les notions « majoré, minoré, croissant, décroissant, monotone » n'ont pas de sens. Seule la convergence peut être définie.

Définition 33.6

La suite (z_n) est **convergente** vers $l \in \mathbb{C}$ quand $n \rightarrow +\infty$ si :

$$\forall \epsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 \quad |z_n - l| \leq \epsilon,$$

où $|z|$ désigne ici le module d'un nombre complexe z .

5. Comparaisons des suites

Définition 33.7

Soient (u_n) et (v_n) des suites à valeurs dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.

- La suite (u_n) est **négligeable** devant la suite (v_n) , ce que nous notons $u_n = o(v_n)$, si

$$\forall \epsilon > 0, \exists n_0 \in \mathbb{N}, \text{ tel que } n \geq n_0 \implies |u_n| \leq \epsilon |v_n|.$$

- La suite (u_n) est **dominée** par la suite (v_n) , ce que nous notons $u_n = O(v_n)$, si

$$\exists M > 0, \exists n_0 \in \mathbb{N}, \text{ tel que } n \geq n_0 \implies |u_n| \leq M |v_n|.$$

- La suite (u_n) est **équivalente** à la suite (v_n) , ce que nous notons $u_n \sim v_n$, si $u_n - v_n = o(v_n)$.

Proposition 33.3

Soient (u_n) et (v_n) des suites à valeurs dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ avec $v_n \neq 0 \forall n \in \mathbb{N}$

- $u_n \sim o(v_n) \iff \lim_{n \rightarrow +\infty} \frac{u_n}{v_n} = 0.$
- $u_n \sim O(v_n) \iff$ la suite $\left(\frac{u_n}{v_n}\right)$ est bornée.
- $u_n \sim v_n \iff \lim_{n \rightarrow +\infty} \frac{u_n}{v_n} = 1.$

Théorème 33.8

Soient (u_n) et (v_n) des suites à valeurs dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, telle que la suite (v_n) est convergente.

- Si $\lim_{n \rightarrow +\infty} v_n = 0$ et si $u_n = o(v_n)$, alors (u_n) converge et $\lim_{n \rightarrow +\infty} u_n = 0$.
- Si $\lim_{n \rightarrow +\infty} v_n = 0$ et si $u_n = O(v_n)$, alors (u_n) converge et $\lim_{n \rightarrow +\infty} u_n = 0$.
- Si $\lim_{n \rightarrow +\infty} v_n = M \in \mathbb{K}$ et si $u_n \sim v_n$, alors (u_n) converge et $\lim_{n \rightarrow +\infty} u_n = M$.

Application

La méthode de Newton

De nombreuses équations rencontrées en physique ne peuvent être résolues que de façon approchée. Parmi les méthodes utilisées, la méthode de Newton permet de résoudre $f(x) = 0$.

Si la fonction f est strictement monotone sur $[a; b]$ et si f vérifie $f(a) \times f(b) < 0$, alors il existe un unique point c de $[a; b]$ tel que $f(c) = 0$. La suite (x_n) définie par $x_0 \in [a; b]$ et

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

converge vers c qui est la solution de $f(x) = 0$.

Une cavité sphérique rigide de centre O et de rayon R contient un fluide homogène. Des ondes sonores stationnaires à symétrie sphérique peuvent s'établir dans la cavité. La vitesse du fluide sur la paroi sphérique est nulle si et seulement si $\tan\left(\frac{wR}{c_s}\right) = \frac{wR}{c_s}$, où w est la pulsation de l'onde et c_s la célérité du son dans le fluide. Calculer, à 10^{-6} près, le plus petit $x > 0$ tel que $\tan(x) = x$.

Posons $f(x) = \tan(x) - x$. f est définie sur $\left]-\frac{\pi}{2} + k\pi; +\frac{\pi}{2} + k\pi\right[$ avec $k \in \mathbb{Z}$.

L'énoncé suppose que $x > 0$, nous restreignons l'intervalle d'étude à $I_1 = \left]0; \frac{\pi}{2}\right[$. Sur I_1 , f est strictement croissante et à valeurs dans $]0; +\infty[$. En effet, f est dérivable sur I et pour tout $x \in I$, $f'(x) = \tan^2(x) > 0$. Donc il n'y a pas de c appartenant à $\left]0; \frac{\pi}{2}\right[$ tel que $f(c) = 0$.

Nous devons rechercher la solution dans un autre intervalle, étudions f sur $I_2 = \left[\frac{\pi}{2}; \frac{3\pi}{2}\right[$. Sur I_2 , f est toujours strictement croissante pour les mêmes raisons et à valeurs dans $]-\infty; +\infty[$. Introduisons la suite (x_n) définie, si $\tan(x_n) \neq 0$, par :

$$\begin{cases} x_0 = \frac{9}{2} \\ x_{n+1} = x_n - \frac{\tan(x_n) - x_n}{\tan^2(x_n)}. \end{cases}$$

Après quatre itérations du processus, nous trouvons $c = 4,493409$.

 Fiche 38

Pour s'entraîner (solutions p. 520)

33.1 Montrer que la suite (u_n) définie par $u_n = n^3 + 7$ est divergente.

33.2 Comparer les suites de terme général $t_n = \frac{1}{n^2}$ et $u_n = \frac{1}{n}$, puis $v_n = \frac{1}{2n}$ et $w_n = \frac{1}{n}$ et enfin $x_n = \frac{1}{n}$ et $y_n = \frac{1}{n-1}$.

33.3 Soient (u_n) définie pour tout $n > 0$ par $u_n = 3 + \frac{1}{n}$ et (v_n) définie pour tout $n > 0$ par $v_n = 3 - \frac{1}{n}$. Montrer que ces deux suites convergent et ont la même limite.

33.4 Reprendre les suites définies dans les Applications 1 et 2 de la Fiche 32 et montrer que ces suites convergent.

33.5 Les suites $(u_n)_{n \in \mathbb{N}} = \left(\frac{2n + (-1)^n \sqrt{n}}{n+1}\right)_{n \in \mathbb{N}}$ et $(v_n)_{n \in \mathbb{N}} = \left(\frac{\sqrt{n}}{\ln(n)}\right)_{n \in \mathbb{N}}$ sont-elles convergentes ?

33.6 Étudier la convergence des suites définies par leur terme général, $u_n = \frac{n^{11}}{10^n}$, $v_n = \sqrt[n]{n} = n^{1/n}$ et $w_n = \frac{n}{n+1} \cos(n\pi)$, et donner leur limite quand c'est possible.

33.7 Montrer que la suite $(u_n)_{n \in \mathbb{N}^*}$ définie par $u_n = 1 + \frac{1}{2^2} + \frac{1}{3^3} + \dots + \frac{1}{n^n}$ est convergente. Vous pouvez utiliser l'égalité :

$$\lim_{n \rightarrow +\infty} \sum_{k=0}^n \frac{1}{2^k} = 2.$$

Suites arithmétiques et géométriques

Dans cette Fiche, nous considérons des suites de nombres réels ou complexes.

1. Suites arithmétiques

Définition 34.1

Une **suite** (u_n) est **arithmétique** de raison $r \in \mathbb{R}$ ou $\mathbb{C}$ si :

$$\forall n \in \mathbb{N}, \quad u_{n+1} = u_n + r.$$

Proposition 34.1 Terme général et somme des premiers termes

- Le terme général d'une suite arithmétique s'écrit :

$$\forall n \in \mathbb{N}, \quad u_n = u_0 + n \times r.$$

- La somme des n premiers termes est égale à :

$$S_n = u_0 + u_1 + \cdots + u_{n-1} = \sum_{k=0}^{n-1} u_k = n \times \left(\frac{u_0 + u_{n-1}}{2} \right).$$

Proposition 34.2 Sens de variation et convergence

- Soit $(u_0, r) \in \mathbb{R}^2$. La monotonie d'une suite arithmétique à valeurs dans $\mathbb{R}$ dépend du signe de la raison r , car, pour tout $n \in \mathbb{N}$, $u_{n+1} - u_n = r$. Si $r > 0$ la suite arithmétique est strictement croissante, si $r < 0$ la suite arithmétique est strictement décroissante et si $r = 0$ la suite arithmétique est constante.
- Soit $(u_0, r) \in \mathbb{R}^2$. Le comportement à l'infini d'une suite arithmétique dépend de la raison r . Si $r = 0$, alors la suite arithmétique est constante et converge vers la constante. Si $r \neq 0$, alors la suite arithmétique diverge soit vers $+\infty$ si $r > 0$ soit vers $-\infty$ si $r < 0$.

2. Suites géométriques

Définition 34.2

Une **suite** (u_n) est **géométrique** de raison $q \in \mathbb{C}$ si :

$$\forall n \in \mathbb{N}, \quad u_{n+1} = q \times u_n.$$

Proposition 34.3 Terme général et somme des premiers termes

- Le terme général d'une suite géométrique s'écrit :

$$\forall n \in \mathbb{N} \quad u_n = u_0 \times q^n.$$

- La somme des n premiers termes est égale à :

$$S_n = \sum_{k=0}^{n-1} u_k = u_0 \times \left(\frac{1 - q^n}{1 - q} \right) \quad \text{si } q \neq 1 \quad \text{et} \quad S_n = n \times u_0 \quad \text{si } q = 1.$$

Conversion : $0^\circ = 1$.

Exemple

Soit la suite géométrique (u_n) de raison $q = \frac{7}{2}$ et de premier terme $u_0 = 1$. Alors la somme des n premiers termes notée S_n est égale à :

$$S_n = -\frac{2}{5} \left(1 - \left(\frac{7}{2} \right)^n \right).$$

Proposition 34.4 Sens de variation et convergence

- Soit $q \in \mathbb{R}$. Le sens de variation d'une suite géométrique à valeurs dans $\mathbb{R}$ dépend du signe de la raison q . Si $u_0 = 0$, alors la suite géométrique est constante et pour tout $n \in \mathbb{N}$, $u_n = 0$. Si $q = 0$, alors la suite géométrique est décroissante si $u_0 \geq 0$ et croissante si $u_0 \leq 0$. Si $u_0 \neq 0$ et $q \neq 0$, alors, pour tout $n \in \mathbb{N}$ $\frac{u_{n+1}}{u_n} = q$. Si $q \in]0; 1[$, alors la suite géométrique est strictement décroissante si $u_0 > 0$ et strictement croissante si $u_0 < 0$. Si $q \in]1; +\infty[$, alors la suite géométrique est strictement croissante si $u_0 > 0$ et strictement décroissante si $u_0 < 0$. Si $q = 1$, alors la suite est constante. Si $q < 0$, alors la suite est alternée.
- Soient $q \in \mathbb{C}$ et $|q|$ son module. Si $|q| < 1$, alors la suite géométrique converge vers 0. Si $|q| > 1$, alors la suite géométrique diverge. Si $q = 1$, alors la suite est constante et converge vers le premier terme. Si $q = -1$, alors la suite diverge.

 Fiche 32

3. Suites arithmético-géométriques**Définition 34.3**

La suite (u_n) à valeurs dans $\mathbb{R}$ est une **suite arithmético-géométrique** quand il existe deux réels a et b tels que la suite (u_n) vérifie la relation de récurrence suivante : $u_{n+1} = au_n + b$.

Proposition 34.5 Terme général et somme des premiers termes

- Si le réel a est différent de 1, alors le terme général de la suite (u_n) est égal à :

$$u_n = a^n \left(u_0 - \frac{b}{1 - a} \right) + \frac{b}{1 - a}.$$

Si $a = 1$, alors le terme général de (u_n) est celui d'une suite arithmétique de raison b .

- Si le réel a est différent de 1, alors la somme des n premiers termes est égale à :

$$S_n = \sum_{k=0}^{n-1} u_k = \left(u_0 - \frac{b}{1-a}\right) \frac{1-a^n}{1-a} + n \frac{b}{1-a}.$$

Si $a = 1$, alors la somme des n premiers termes est égale à la somme d'une suite arithmétique.

Proposition 34.6

Le terme général et les considérations sur les suites géométriques permettent de déterminer la limite d'une telle suite suivant les valeurs de a et le signe de $u_n - \frac{b}{1-a}$ si $a \neq 1$.



Une remarque intéressante est à faire dans le cas où $|a| < 1$. Dans ce cas, la limite de la suite est égale à $\frac{b}{1-a}$ et ce quelle que soit la valeur initiale.

Application L'interféromètre à réseau

Nous étudions un interféromètre à réseau constitué de N radiotélescopes. Nous obtenons comme amplitude complexe du signal :

$$\underline{a}(\phi) = \underline{a}_0 \sum_{k=1}^N e^{i(k-1)\phi}$$

où $\underline{a}_0 \in \mathbb{C}$ désigne l'amplitude et $\phi \in [0; 2\pi]$ la phase. Calculer $\underline{a}(\phi)$.

En développant l'expression, nous obtenons :

$$\underline{a}(\phi) = \underline{a}_0 (1 + e^{i\phi} + e^{i2\phi} + \dots + e^{i(N-1)\phi}) = \underline{a}_0 (e^{i0\phi} + e^{i\phi} + e^{i2\phi} + \dots + e^{i(N-1)\phi}).$$



Nous reconnaissons la somme de N termes d'une suite géométrique de raison $e^{i\phi}$. Reste à distinguer deux cas : soit $e^{i\phi} = 1$, soit $e^{i\phi} \neq 1$.

Commençons par le cas où $e^{i\phi} = 1$. Alors $\phi = 0 \pmod{2\pi}$. Dans ce cas, $\underline{a}(\phi) = N\underline{a}_0$. Maintenant, traitons le cas où $e^{i\phi} \neq 1$. Alors, nous avons :

$$\underline{a}(\phi) = \underline{a}_0 \frac{1 - e^{iN\phi}}{1 - e^{i\phi}}.$$



Nous allons factoriser au numérateur et au dénominateur par une quantité adéquate : cette méthode est à retenir !

Nous obtenons :

$$\underline{a}(\phi) = \underline{a}_0 \frac{e^{\frac{iN\phi}{2}} \left(e^{-\frac{iN\phi}{2}} - e^{\frac{iN\phi}{2}} \right)}{e^{\frac{i\phi}{2}} \left(e^{-\frac{i\phi}{2}} - e^{\frac{i\phi}{2}} \right)}.$$

En appliquant les formules d'Euler, nous en déduisons :

$$\underline{a}(\phi) = \underline{a}_0 e^{\frac{i(N-1)\phi}{2}} \frac{\sin\left(\frac{N\phi}{2}\right)}{\sin\left(\frac{\phi}{2}\right)}.$$



Fiche 16

Pour s'entraîner (solutions p. 521)

34.1 Soit la suite (u_n) définie par $u_{n+1} = 5u_n$ et $u_0 = 2$. Montrer que la suite (u_n) diverge.

34.2 Une suite arithmético-géométrique

Nous considérons la suite numérique $(u_n)_{n \in \mathbb{N}}$ définie, pour tout $n \in \mathbb{N}$, par :
$$\begin{cases} u_0 = \frac{3}{5} \\ u_{n+1} = \frac{u_n - 3}{6} \end{cases}$$

Soit $(v_n)_{n \in \mathbb{N}}$ la suite définie par : $\forall n \in \mathbb{N}, v_n = 5u_n + 3$.

1. Démontrer que (v_n) est une suite géométrique. En déduire une expression de u_n en fonction de n .

2. Démontrer que la suite (u_n) est décroissante. Est-elle convergente ?

3. Calculer $S_n = v_0 + v_1 + \dots + v_{n-1}$ et $S'_n = u_0 + u_1 + \dots + u_{n-1}$.
Déterminer les limites de S_n et S'_n quand $n \rightarrow +\infty$

34.3 Soient la suite (u_n) définie, pour tout $n \in \mathbb{N}$ par $u_0 = 1$ et $u_{n+1} = \frac{1}{3}u_n + n - 1$ et la suite (v_n) définie, pour tout $n \in \mathbb{N}$ par $v_n = 4u_n - 6n + 15$.

1. Montrer que (v_n) est une suite géométrique. Exprimer le terme général v_n en fonction de n .

2. En déduire que : $\forall n \in \mathbb{N}, u_n = \frac{19}{4 \times 3^n} + \frac{6n - 15}{4}$.

3. Montrer que le terme général de la suite (u_n) peut s'écrire sous la forme $u_n = t_n + w_n$ où t_n est le terme d'une suite géométrique et w_n est le terme d'une suite arithmétique.

4. Calculer les trois sommes suivantes : $T_n = t_0 + t_1 + \dots + t_n$, $W_n = w_0 + w_1 + \dots + w_n$, $U_n = u_0 + u_1 + \dots + u_n$.

34.4 Soient les deux suites (u_n) et (v_n) définies pour tout $n \in \mathbb{N}$ par :

$$\begin{aligned} u_0 &= 3, \quad u_{n+1} = \frac{u_n + v_n}{2}; \\ v_0 &= 4, \quad v_{n+1} = \frac{u_{n+1} + v_n}{2}. \end{aligned}$$

1. Calculer u_1, v_1, u_2 et v_2 .

2. Montrer que (w_n) définie, pour tout $n \in \mathbb{N}$, par $w_n = v_n - u_n$, est une suite géométrique de raison $\frac{1}{4}$. Exprimer w_n en fonction de n .

3. Après avoir étudié le sens de la variation des suites (u_n) et (v_n) , démontrer que ces deux suites convergent vers une même limite que nous noterons L .

4. Montrer que la suite (j_n) définie, pour tout $n \in \mathbb{N}$, par $j_n = \frac{u_n + 2v_n}{3}$, est constante. En déduire la limite des suites (u_n) et (v_n) .

1. Suite récurrente linéaire d'ordre deux

Définition 35.1

Une **suite récurrente linéaire d'ordre deux** est une suite (u_n) dont le terme général satisfait la relation de récurrence d'ordre deux suivante :

$$\forall n \in \mathbb{N}, \quad u_{n+2} = au_{n+1} + bu_n \quad (35.1)$$

où a et b sont des nombres réels ou complexes, avec $b \neq 0$. Les deux premiers termes de la suite, à savoir u_0 et u_1 , sont fixés dans $\mathbb{R}$ ou dans $\mathbb{C}$.

L'ensemble E des suites $(u_n)_{n \in \mathbb{N}}$ dont le terme général satisfait une relation de la forme $u_{n+2} = au_{n+1} + bu_n$ est un espace vectoriel de dimension 2. Nous en cherchons une base en résolvant l'équation caractéristique suivante :

 **Fiche 18**

$$r^2 - ar - b = 0.$$

► Premier cas : a et b deux complexes

- Si le discriminant, noté Δ , est non nul, l'équation caractéristique admet deux solutions distinctes notées r_1 et r_2 . Toute suite vérifiant la relation (35.1) est alors du type :

$$u_n = C_1 r_1^n + C_2 r_2^n$$

où C_1 et C_2 sont deux constantes complexes que nous déterminerons à l'aide de u_0 et de u_1 .

- Si le discriminant, noté Δ , est nul, l'équation caractéristique admet une solution double notée r_0 . Toute suite vérifiant la relation (35.1) est alors du type :

$$u_n = (C_1 + C_2 n) r_0^n$$

où C_1 et C_2 sont deux constantes complexes que nous déterminerons à l'aide de u_0 et de u_1 .

► Second cas : a et b deux réels

- Si le discriminant, noté Δ , est positif ou nul, la suite (u_n) est du même type que précédemment, les constantes C_1 et C_2 sont ici réelles.
- Si le discriminant, noté Δ , est strictement négatif, l'équation caractéristique admet deux solutions complexes conjuguées notées r_1 et r_2 , que nous pouvons écrire sous forme trigonométrique, $r_1 = \rho e^{i\theta}$ et $r_2 = \rho e^{-i\theta}$, avec $\rho > 0$. Toute suite vérifiant la relation (35.1) est alors du type :

$$u_n = \rho^n (C_1 \cos(n\theta) + C_2 \sin(n\theta))$$

où C_1 et C_2 sont deux constantes réelles que nous déterminerons à l'aide de u_0 et de u_1 .

2. Suites récurrentes du type $u_{n+1} = f(u_n)$

Nous considérons ici une fonction f à valeurs réelles.

► Vocabulaire

Définition 35.2

- Un **intervalle** I est **stable** pour la fonction f lorsque l'intervalle I est inclus dans l'espace de départ de la fonction f et que $f(I) \subset I$, i.e. $\forall x \in I, f(x) \in I$.
- Le point a appartenant à l'espace de départ de la fonction f est un **point fixe** de f lorsque $f(a) = a$.

Proposition 35.1

Lorsque I est un intervalle stable pour la fonction f , pour toute condition initiale $u_0 \in I$, la suite (u_n) de terme général $u_{n+1} = f(u_n)$ est définie pour tous les entiers n car $u_n \in I$ pour tout $n \in \mathbb{N}$.

Définition 35.3

- Un **point fixe** a de la fonction f est **stable ou attractif** si et seulement s'il existe un intervalle J stable pour f tel que a est à l'intérieur de J et que, pour toute condition initiale $u_0 \in J$, la suite (u_n) converge vers a .
- Un **point fixe** a de la fonction f est **instable ou répulsif** si et seulement s'il existe un intervalle J tel que a est à l'intérieur de J et que, pour toute condition initiale $u_0 \in J$, il existe un entier $N(u_0)$ tel que :

$$u_0 \in J; u_1 \in J; \dots; u_{N(u_0)-1} \in J; u_{N(u_0)} \notin J.$$

► Étude pratique d'une suite de type $u_{n+1} = f(u_n)$

Première étape

Pour étudier une telle suite, nous déterminons d'abord un intervalle contenant toutes les valeurs de la suite. En fait, nous recherchons les intervalles stables par la fonction f . Cette étude revient à chercher un intervalle où soient situés tous les termes de la suite.

Deuxième étape : les limites éventuelles

Ensuite, nous cherchons la limite éventuelle de la suite. Deux cas se présentent alors :

1. soit la suite est divergente, et dans ce cas l'étude est achevée ;
2. soit la suite est convergente. Dans ce cas, nous cherchons les points fixes à savoir les solutions de l'équation $f(x) = x$. Les solutions de cette équation sont les éventuels candidats pour être la limite de la suite. En effet, nous avons le résultat suivant :

Théorème 35.1

Soit $(u_n)_{n \in \mathbb{N}}$ une suite récurrente du type $u_{n+1} = f(u_n)$. Si la suite converge vers l et si la fonction f est continue en l , alors l est un point fixe de f . Autrement dit, la limite l est solution de l'équation $f(x) = x$.



En général, la fonction f possède non pas un mais plusieurs points fixes. Pour déterminer la limite éventuelle, nous utilisons le résultat classique sur les suites : si, $\forall n \in \mathbb{N}, u_n \in]a; b[$ et si la suite (u_n) converge vers l alors $l \in [a; b]$.

Troisième étape : la monotonie de la suite

1. Si la fonction f est croissante sur l'intervalle stable, alors la suite (u_n) est monotone. Il ne reste plus qu'à étudier le signe de $u_1 - u_0$ pour savoir si la suite (u_n) est croissante ou décroissante. Si $f(x) - x \geq 0$ pour tout $x \in I$ un intervalle stable de f contient les $u_n, n \in \mathbb{N}$ alors la suite est croissante.
2. Si la fonction f est décroissante sur l'intervalle stable, alors les suites (u_{2n}) et (u_{2n+1}) sont monotones et de sens contraire.

Application

1. Soit la suite (u_n) la suite récurrente d'ordre deux définie par : $u_0 = 3, u_1 = -6$ et pour tout $n \in \mathbb{N}, u_{n+2} = 4u_{n+1} - 16u_n$. Déterminer le terme général u_n de la suite (u_n) .

2. Un homme place un couple de lapins dans une très grande cage. Combien de couples de lapins obtiendra-t-il après 12 mois si chaque couple engendre tous les mois un et un seul nouveau couple à compter du second mois de son existence ?

1. Nous introduisons l'équation caractéristique : $r^2 - 4r + 16 = 0$. Nous calculons le discriminant : $\Delta = -48 = (4\sqrt{3}i)^2 < 0$. Donc l'équation caractéristique a deux solutions complexes conjuguées qui sont égales respectivement à $r_1 = 2 + 2\sqrt{3}i$ et $r_2 = 2 - 2\sqrt{3}i$. Mettons ces deux solutions sous forme d'exponentielle complexe.

$$r_1 = 4 \left(\frac{1}{2} + i \frac{\sqrt{3}}{2} \right) = 4 \left(\cos \left(\frac{\pi}{3} \right) + i \sin \left(\frac{\pi}{3} \right) \right) = 4e^{i\frac{\pi}{3}}$$

et

$$r_2 = 4 \left(\frac{1}{2} - i \frac{\sqrt{3}}{2} \right) = 4 \left(\cos \left(\frac{\pi}{3} \right) - i \sin \left(\frac{\pi}{3} \right) \right) = 4e^{-i\frac{\pi}{3}}.$$

Puis nous avons

$$u_n = 4^n \left(C_1 \cos \left(\frac{n\pi}{3} \right) + C_2 \sin \left(\frac{n\pi}{3} \right) \right) \quad \text{où } C_1 \text{ et } C_2 \text{ sont deux constantes réelles.}$$

Comme $u_0 = 3$ et $u_1 = -6$, nous en déduisons les valeurs de C_1 et de C_2 . Après calculs, nous obtenons : $C_1 = 3$ et $C_2 = -2\sqrt{3}$. Donc, en remplaçant C_1 et C_2 par leurs valeurs respectives et après simplifications, nous obtenons :

$$u_n = 4^n \sqrt{3} \left(\sqrt{3} \cos \left(\frac{n\pi}{3} \right) - 2 \sin \left(\frac{n\pi}{3} \right) \right).$$

2. Soit F_n le nombre de couples de lapins au début du mois n . Jusqu'à la fin du deuxième mois, la population se limite au premier couple : $F_1 = F_2 = 1$. Au mois $n + 2$, le nombre de couples F_{n+2} est la somme du nombre F_{n+1} de ceux qui existaient le mois précédent et du nombre de couples nouvellement engendrés, donc $F_{n+2} = F_{n+1} + F_n$.

L'équation caractéristique est $r^2 - r - 1 = 0$. Les deux solutions de cette équation sont $r_1 = \frac{1 - \sqrt{5}}{2}$ et $r_2 = \frac{1 + \sqrt{5}}{2}$. Ainsi, pour tout $n \geq 1$, $F_n = C_1 \left(\frac{1 - \sqrt{5}}{2} \right)^n + C_2 \left(\frac{1 + \sqrt{5}}{2} \right)^n$.

Nous calculons maintenant les deux constantes C_1 et C_2 . Nous avons d'une part $F_0 = 0$ et par conséquent $C_1 + C_2 = 0$. D'autre part nous avons $F_1 = 1$ et par conséquent $C_1 \left(\frac{1 - \sqrt{5}}{2} \right) + C_2 \left(\frac{1 + \sqrt{5}}{2} \right) = 1$. Donc, après la résolution d'un système d'équations à deux inconnues, nous obtenons $C_1 = -\frac{1}{\sqrt{5}}$ et $C_2 = \frac{1}{\sqrt{5}}$. Donc, pour tout $n \geq 1$, $F_n = \frac{1}{\sqrt{5}} \left(\left(\frac{1 + \sqrt{5}}{2} \right)^n - \left(\frac{1 - \sqrt{5}}{2} \right)^n \right)$.



Le nombre réel $\frac{1 + \sqrt{5}}{2}$ est appelé le nombre d'or et la suite (F_n) s'appelle la suite de Fibonacci.

Pour s'entraîner (solutions p. 522)

35.1 Étudier la suite définie pour tout $n \in \mathbb{N}^*$, par : $u_1 = -\frac{1}{2}$ et $u_{n+1} = u_n^2 + \frac{1}{4}$.

35.2 Soient $I = [0; 1]$ et f la fonction définie sur I par : $f(x) = \frac{3x+2}{x+4}$.

1. Donner le tableau de variation de f . En déduire que, pour tout $x \in I$, $f(x)$ appartient à I .

2. Soit la suite (u_n) définie par : $u_0 = 0$ et $u_{n+1} = \frac{3u_n+2}{u_n+4}$. Montrer que pour tout $n \in \mathbb{N}$, $u_n \in I$.

3. Étudier le sens de variation de la suite (u_n) .

4. Démontrer que la suite (u_n) est convergente.

5. Prouver que la limite l de la suite (u_n) vérifie $l = f(l)$ et calculer l .

35.3 

Soient $A \in \mathbb{R}$ et f la fonction définie sur $\mathbb{R}$ par $f(x) = x^2 + A$.

1.a) Donner le tableau de variation de f .

1.b) Donner le tableau de signe de $f(x) - x$ selon la valeur de A par rapport à $1/4$.

Soit la suite (u_n) définie par : $u_0 = 0$ et $u_{n+1} = u_n^2 + A$.

1.c) Montrer que si I est un intervalle stable par f et si $u_0 \in I$, alors pour tout $n \in \mathbb{N}$, $u_n \in I$.

Dans cette question, nous supposons $A \geq 0$.

2.a) Montrer que (u_n) est croissante.

2.b) Montrer que si $A > 1/4$, alors (u_n) tend vers $+\infty$.

2.c) Montrer que si $A \in [0; 1/4[$, alors (u_n) converge et donner sa limite.

Nous supposons dans cette question que $A \in]-1; 0[$.

3.a) Montrer que $[A; 0]$ est stable par f .

3.b) Montrer que (u_{2n}) est décroissante et converge vers $a \in [A; 0]$ tel que $f \circ f(a) = a$. Montrer que (u_{2n+1}) est croissante et converge vers $b \in [A; 0]$ tel que $f \circ f(b) = b$.

3.c) Montrer que pour tout $x \in \mathbb{R}$, $f \circ f(x) - x = (x^2 - x + A) \times (x^2 + x + A + 1)$.

3.d) Montrer que si $A \in]-3/4; 0[$, alors (u_{2n}) et (u_{2n+1}) convergent vers la même limite. En déduire que (u_n) converge et donner sa limite.

3.e) Montrer que si $A \in]-1; -3/4[$, alors (u_n) diverge.

Continuité d'une fonction

Nous considérons des fonctions $f : I \longrightarrow \mathbb{R}$ définies sur un intervalle I de $\mathbb{R}$ et à valeurs réelles. Commençons par définir la notion de limite d'une fonction en un point.

Définition 36.1

Soient $f : I \longrightarrow \mathbb{R}$, $x_0 \in \mathbb{R}$ et $\ell \in \mathbb{R}$. Nous disons que la **fonction** f admet la limite ℓ au point x_0 , et nous notons $\lim_{x \rightarrow x_0} f(x) = \ell$ si

$$\forall \varepsilon > 0, \exists \eta > 0, \forall x \in I, \quad (|x - x_0| \leq \eta \implies |f(x) - \ell| \leq \varepsilon).$$

Théorème 36.1

Soient deux fonctions f et g définies sur un intervalle I de $\mathbb{R}$, à valeurs réelles. Soit $x_0 \in \mathbb{R}$ tel que $\lim_{x \rightarrow x_0} f(x) = \ell_1$ et $\lim_{x \rightarrow x_0} g(x) = \ell_2$ alors $\lim_{x \rightarrow x_0} (f + g)(x) = \ell_1 + \ell_2$; $\lim_{x \rightarrow x_0} (f \times g)(x) = \ell_1 \ell_2$, et si $\ell_2 \neq 0$, $\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = \frac{\ell_1}{\ell_2}$.

1. Continuité en un point et sur un intervalle

Définition 36.2

Soient $f : I \longrightarrow \mathbb{R}$ et $x_0 \in I$. La **fonction** f est **continue** au point x_0 si la fonction f est définie en x_0 et si $\lim_{x \rightarrow x_0} f(x) = f(x_0)$ ce qui s'écrit mathématiquement :

$$\forall \varepsilon > 0, \exists \eta > 0, \forall x \in I, \quad (|x - x_0| \leq \eta \implies |f(x) - f(x_0)| \leq \varepsilon).$$

La **fonction** f est **discontinue** au point x_0 si et seulement si f n'est pas continue en x_0 .

Proposition 36.1

Soient $f : I \longrightarrow \mathbb{R}$ et $x_0 \in I$. Pour que f soit continue en x_0 il faut et il suffit que f admette $f(x_0)$ pour limite en x_0 .

Définition 36.3

La **fonction** f est **continue à gauche** en x_0 (respectivement à **droite** en x_0) si la fonction f est définie en x_0 et si $\lim_{x \rightarrow x_0^-} f(x) = f(x_0)$ (respectivement $\lim_{x \rightarrow x_0^+} f(x) = f(x_0)$).

Définition 36.4 Prolongement par continuité

Soient f une fonction définie sur un intervalle I et $x_0 \notin I$. Si nous avons $\lim_{x \rightarrow x_0} f(x) = l$, avec $l \in \mathbb{R}$, la fonction $\tilde{f}$ définie sur $I \cup \{x_0\}$ par $\tilde{f}(x_0) = l$ et $\tilde{f}(x) = f(x)$ pour $x \in I$, est la seule fonction continue au point x_0 dont la restriction à I soit f . Cette dernière est appelée le **prolongement par continuité** de la fonction f au point x_0 .

Définition 36.5

Soit E un ensemble qui soit un intervalle ou une réunion d'intervalles. Une **fonction** f , définie sur E est **continue sur** E si f est continue en tout point de E .

Théorème 36.2

- Soient $\lambda \in \mathbb{R}$ et I un intervalle. Si les fonctions f et g sont définies et continues en $x_0 \in I$, alors $f + g$, $f \times g$, λf et $\frac{f}{g}$, avec la condition que $g(x_0) \neq 0$ sont continues en x_0 .
- **Fonction composée** : Si f est définie et continue en x_0 et si g est définie et continue en $f(x_0)$, alors la fonction $g \circ f$ est définie et continue en x_0 .



En fait, l'ensemble noté $\mathcal{C}(I)$ des fonctions continues sur I constitue un anneau et un $\mathbb{R}$ -espace vectoriel pour les lois usuelles $+$ et $\times$.

Fiches 8 et 18

2. Image d'un intervalle**Définition 36.6**

Soient I un intervalle et f une fonction définie sur I . L'ensemble des images par f des éléments de I est appelé **image de I par f** et est noté $f(I)$. Nous avons $f(I) = \{y \in \mathbb{R}, \exists x \in I, y = f(x)\}$.

Théorème 36.3

Soit f une fonction définie et continue sur un intervalle I et à valeurs réelles, alors l'image $f(I)$ est un intervalle.

Voici un énoncé équivalent mais plus explicite :

Théorème 36.4 Théorème des valeurs intermédiaires

Soit f une fonction définie et continue sur un intervalle $[a; b]$ et à valeurs réelles. Alors $\forall u \in [f(a); f(b)]$, il existe au moins un élément c compris entre a et b tel que $u = f(c)$.

Théorème 36.5

Théorème de Bolzano Soit f une fonction définie et continue sur $[a; b]$ et à valeurs réelles. Si $f(a) \times f(b) < 0$, alors il existe au moins un réel c compris entre a et b tel que $f(c) = 0$.

Proposition 36.2

Pour tout polynôme P à coefficients réels et de degré impair, il existe au moins une racine réelle, c'est-à-dire un réel c tel que $P(c) = 0$.

Théorème 36.6

L'image d'un intervalle fermé borné par une fonction continue est un intervalle fermé borné.

Définition 36.7

- Une fonction f définie sur un intervalle I à valeurs réelles est dite **majorée** si et seulement s'il existe $M \in \mathbb{R}$ tel que : $\forall x \in I, f(x) \leq M$.
- Une fonction f définie sur un intervalle I à valeurs réelles est dite **minorée** si et seulement s'il existe $m \in \mathbb{R}$ tel que : $\forall x \in I, m \leq f(x)$.
- Une fonction f définie sur un intervalle I à valeurs réelles est dite **bornée** si et seulement s'il existe $m, M \in \mathbb{R}$ tels que : $\forall x \in I, m \leq f(x) \leq M$.

Proposition 36.3

Si f est une fonction continue sur $[a; b]$, alors f est bornée et atteint ses bornes sur $[a; b]$.



Attention, une fonction continue sur un intervalle I non fermé ou non continue sur I peut ne pas être bornée sur I . Par exemple, la fonction $f : x \mapsto \frac{1}{x}$ est continue sur $]0; 1]$ (intervalle non fermé) mais f n'est pas majorée sur $]0; 1]$.

3. Fonction réciproque

Théorème 36.7

Soit f une fonction continue et strictement croissante (resp. décroissante) sur un intervalle I . Alors f est une bijection de I sur l'intervalle $f(I)$ et sa bijection réciproque notée f^{-1} est continue et strictement croissante (resp. décroissante) sur l'intervalle $f(I)$.



Dans un plan rapporté à un repère orthonormé, les courbes représentatives de f et f^{-1} sont symétriques par rapport à la première bissectrice.

Exemples

1. La fonction logarithme népérien $\ln$ est continue et strictement croissante sur $]0; +\infty[$. Donc la fonction $\ln$ est une bijection de $]0; +\infty[$ sur $\mathbb{R}$ car $\lim_{x \rightarrow 0^+} \ln(x) = -\infty$ et $\lim_{x \rightarrow +\infty} \ln(x) = +\infty$. La bijection réciproque est la fonction exponentielle, notée $\exp$.
2. Soit la fonction $f(x) = x^n$ avec $n \geq 2$. La fonction f est continue sur $\mathbb{R}$ et strictement croissante sur $[0; +\infty[$. Sa restriction à $[0; +\infty[$ est une bijection de $[0; +\infty[$ sur $[0; +\infty[$. La bijection réciproque est la fonction $f^{-1}(x) = (x)^{1/n}$.

Application

1. Étudier la continuité de la fonction f définie sur $\mathbb{R}$ par : $f(x) = \frac{x}{1 + |x|}$. Puis montrer que f admet une fonction réciproque f^{-1} .

2. Application du théorème des valeurs intermédiaires

Soit f une application continue de $[0; 1]$ dans $\mathbb{R}$ telle que : $f(0) = f(1)$.

a. Montrer que pour tout $n \in \mathbb{N}^*$, il existe un réel x_n dans $\left[0; 1 - \frac{1}{n}\right]$ tel que :

$$f(x_n) = f\left(x_n + \frac{1}{n}\right).$$

b. Montrer que si une voiture parcourt 100 kilomètres en une heure, il existe alors un intervalle de temps égal à une demi-heure pendant lequel la voiture a parcouru 50 kilomètres.

1. La fonction f est définie sur $\mathbb{R}$. La fonction $x \mapsto 1 + |x|$ est continue sur $\mathbb{R}$ et ne s'annule pas sur $\mathbb{R}$ donc f est continue sur $\mathbb{R}$. $\forall x \in \mathbb{R}$, $f(-x) = -f(x)$ donc f est impaire sur $\mathbb{R}$. $\forall x > 0$, $f(x) = \frac{x}{1+x} = 1 - \frac{1}{1+x}$. La fonction $x \mapsto 1 + x$ est strictement croissante sur $\mathbb{R}_+$ donc la fonction $x \mapsto \frac{1}{1+x}$ est strictement décroissante sur $\mathbb{R}_+$ et f est strictement croissante sur $\mathbb{R}_+$. Étant impaire et continue sur $\mathbb{R}$, f est strictement croissante sur $\mathbb{R}$. f est une bijection de $\mathbb{R}$ sur $] -1; +1[$ donc f a une fonction réciproque $f^{-1} :] -1; +1[\rightarrow \mathbb{R}$, continue et strictement croissante sur $] -1; +1[$.

2.a. Pour tout $n \geq 1$, nous désignons par f_n la fonction définie sur $I_n = \left[0; 1 - \frac{1}{n}\right]$ par $f_n(x) = f\left(x + \frac{1}{n}\right) - f(x)$. Si la fonction f_n ne s'annule jamais sur I_n , du fait de sa continuité, nous déduisons du théorème des valeurs intermédiaires qu'elle garde un signe constant sur cet intervalle. Supposons que pour tout $x \in I_n$, $f_n(x) > 0$. Alors nous avons :

$$f(1) - f(0) = \sum_{k=0}^{n-1} \left(f\left(\frac{k+1}{n}\right) - f\left(\frac{k}{n}\right) \right) = \sum_{k=0}^{n-1} f_n\left(\frac{k}{n}\right) > 0$$

ce qui contredit $f(0) = f(1)$, la fonction f_n s'annule donc au moins une fois sur I_n .

b. Nous considérons la fonction f définie sur $[0; 1]$ par : $f(x) = d(x) - 100x$ où $d(x)$ est le nombre de kilomètres parcourus en x heures. Cette fonction est continue et $f(0) = f(1) = 0$. Le résultat précédent, pour $n = 2$, dit qu'il existe $x_2 \in \left[0; \frac{1}{2}\right]$ tel que $f(x_2) = f\left(x_2 + \frac{1}{2}\right)$, soit $d\left(x_2 + \frac{1}{2}\right) - d(x_2) = 50$ et $\left[x_2; x_2 + \frac{1}{2}\right]$ répond à la question.

Pour s'entraîner (solutions p. 522)

36.1 Étudier la continuité de la fonction f définie sur $\mathbb{R}$ par : $f(x) = E(x) + E(2-x)$.

36.2 Étudier la continuité de la fonction f définie sur $\mathbb{R}$ par : $f(x) = E(x) + \sqrt{x - E(x)}$ où $E(x)$ désigne la partie entière de x .

36.3 Étudier la continuité de la fonction f définie sur $\mathbb{R}$ par : $f(x) = E(x) + (x - E(x))^2$.

36.4 Montrer que la fonction f définie par : $f(x) = 1 - \sqrt{1 - x^2}$ est continue sur l'intervalle $[-1; +1]$.

36.5 Étudier la continuité de la fonction f définie sur $\mathbb{R}$ par : $f(1) = f(-1) = 0$ et pour

tout $x \in \mathbb{R} \setminus \{-1; +1\}$,

$$f(x) = \lim_{n \rightarrow +\infty} \left(\frac{x^{2n} + 1}{x^{2n} - 1} \right).$$

36.6 Montrer que la fonction f définie par

$$f(x) = \sin(x+1) \ln|1+x|$$

est prolongeable par continuité en $x = -1$.

36.7 **1.** Montrer que le fonction f définie par : $f(x) = \sin(x) \sin\left(\frac{1}{x}\right)$ est prolongeable par continuité sur $\mathbb{R}$.

2. Même question avec la fonction f définie par : $f(x) = \cos(x) \cos\left(\frac{1}{x}\right)$.

Dérivabilité d'une fonction

Nous considérons des fonctions à valeurs réelles.

1. Dérivée en un point. Dérivée sur un intervalle

Définition 37.1

Soient f une fonction définie sur un intervalle I et un point $x_0 \in I$ tel que f soit définie au voisinage de x_0 .

1. La fonction f est **dérivable au point** x_0 si et seulement si

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

existe et est finie. Cette limite est un réel noté $f'(x_0)$ ou $\frac{df}{dx}(x_0)$ et est appelée **dérivée de f en x_0** .

2. La fonction f est **dérivable à droite au point** x_0 si et seulement si

$$\lim_{x \rightarrow x_0^+} \frac{f(x) - f(x_0)}{x - x_0}$$

existe et est finie. Cette limite est un réel noté $f'_d(x_0)$ et est appelée **dérivée à droite de f en x_0** .

3. De même, nous pouvons définir la notion de **dérivable à gauche au point** x_0 et de **dérivée à gauche de f en x_0** . Cette limite sera notée $f'_g(x_0)$.

4. La fonction f est **dérivable au point** x_0 si et seulement si la fonction f admet au point x_0 une dérivée à droite et une dérivée à gauche qui sont égales.



Une définition équivalente de $f'(x_0)$ s'obtient en posant $x = x_0 + h$:

$$f'(x_0) = \lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h}.$$

Le rapport $\frac{f(x_0 + h) - f(x_0)}{h}$ s'appelle le **taux d'accroissement** ou **taux de variation de f entre x_0 et $x_0 + h$** .

Définition 37.2

La fonction f est **dérivable sur un intervalle** I si la fonction f est dérivable en tout point de cet intervalle.

Définition 37.3

Soit f une fonction définie sur I à valeurs dans $\mathbb{R}$. Nous appelons **fonction dérivée de f** l'application qui, à chaque $x \in I$ tel que $f'(x)$ existe, associe $f'(x)$. La fonction dérivée de f est notée f' ou $\frac{df}{dx}$.

2. Lien entre dérivabilité et continuité**Théorème 37.1**

Soit une fonction f définie sur I à valeurs dans $\mathbb{R}$ et $x_0 \in I$. Si la fonction f est dérivable en x_0 , alors la fonction f est continue en x_0 .



Attention, une fonction peut être continue en x_0 sans être dérivable en x_0 . En voici un contre-exemple : la fonction valeur absolue est continue en 0 mais n'est pas dérivable en 0.

3. Équation d'une tangente

Si la fonction f est dérivable au point x_0 , cela signifie que la représentation graphique de la fonction f admet au point d'abscisse x_0 une tangente de coefficient directeur $f'(x_0)$.

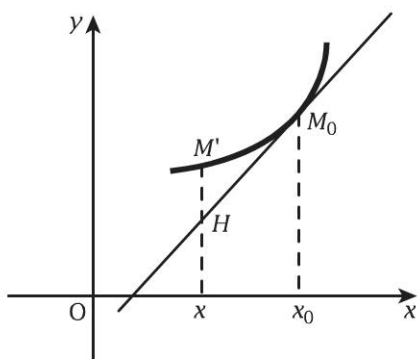
Définition 37.4

L'équation de la **tangente** de la fonction f au point x_0 est égale à :

$$y = f'(x_0) \times (x - x_0) + f(x_0).$$



Si $\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = \pm\infty$ alors la fonction f n'est pas dérivable au point x_0 mais la représentation graphique de f admet au point d'abscisse x_0 une tangente verticale parallèle à l'axe Oy .

**4. Dérivées successives****Définition 37.5**

Soit une fonction f dérivable sur un intervalle I , de fonction dérivée notée f' . Si la fonction f' est dérivable sur l'intervalle I , alors sa fonction dérivée est notée f'' ou encore $f^{(2)}$. De même, nous pouvons définir, par récurrence sur l'entier n , la **dérivée $n^{\text{ème}}$ de f** ou **dérivée d'ordre n de f** sur un intervalle I en posant :

$$f^{(0)} = f, \quad f^{(n)} = (f^{(n-1)})', \quad \text{si } f^{(n-1)} \text{ est dérivable sur } I.$$

Nous disons que f est **n fois dérivable sur I** si et seulement si $f^{(n)}$ est définie sur I . Nous disons que f est **indéfiniment dérivable sur I** si et seulement si f est n fois dérivable sur I pour tout $n \in \mathbb{N}$.

Définition 37.6

Soit f une fonction définie sur un intervalle I à valeurs réelles.

- Soit $n \in \mathbb{N}$. La fonction f est de **classe $\mathcal{C}^n$ sur I** si et seulement si f est n fois dérivable sur I et si la fonction $f^{(n)}$ est continue sur I .
- La fonction f est de **classe $\mathcal{C}^\infty$ sur I** si et seulement si f est indéfiniment dérivable sur I .

5. Opérations sur les fonctions dérivables

► Opérations algébriques des fonctions dérivables en un point

Soient f et g deux fonctions dérivables au point x_0 et un réel λ . Les fonctions $f + g$, λf , $f \times g$, et $\frac{f}{g}$, avec la condition que $g(x_0) \neq 0$, sont dérivables au point x_0 et nous avons alors :

$$\begin{aligned}(f + g)'(x_0) &= f'(x_0) + g'(x_0) \\ (\lambda f)'(x_0) &= \lambda f'(x_0) \\ (f \times g)'(x_0) &= f'(x_0) \times g(x_0) + f(x_0) \times g'(x_0) \\ \left(\frac{f}{g}\right)'(x_0) &= \frac{f'(x_0) \times g(x_0) - f(x_0) \times g'(x_0)}{g^2(x_0)}.\end{aligned}$$

► Dérivée d'une fonction composée

Soit f une fonction dérivable au point x_0 et une fonction g dérivable au point $f(x_0)$, alors la fonction $g \circ f$ est dérivable au point x_0 et nous avons :

$$(g \circ f)'(x_0) = f'(x_0) \times g'(f(x_0)).$$

► Dérivée d'une fonction réciproque

Soit f une fonction continue et strictement monotone sur un intervalle I . Nous supposons que la fonction f est dérivable au point x_0 et que $f'(x_0) \neq 0$. Alors la fonction réciproque, notée f^{-1} est dérivable en $f(x_0)$ et nous avons :

$$(f^{-1})'(f(x_0)) = \frac{1}{f'(x_0)}.$$

► Formule de Leibniz

Si les deux fonctions f et g admettent des dérivées d'ordre n au point x_0 alors il en est de même de la fonction $f \times g$ et nous avons :

$$(f \times g)^{(n)}(x_0) = \sum_{k=0}^n \binom{n}{k} f^{(k)}(x_0) \times g^{(n-k)}(x_0).$$

La résistance R d'une thermistance varie avec la température Kelvin T selon :

$$R(T) = R_1 e^{\left[a\left(\frac{1}{T} - \frac{1}{T_1}\right)\right]},$$

avec $a = 4 \cdot 10^3 \text{ K}$ et $R_1 = 10^3 \Omega$ à $T_1 = 300 \text{ K}$.

1. Déterminer le coefficient de température $k(T) = \frac{1}{R(T)} \frac{dR}{dT}(T)$ et le calculer à 300 K .
2. Sachant que nous mesurons la résistance avec une précision de 0,1%, quelle variation de température pouvons-nous détecter au voisinage de 300 K ?

1. $R(T)$ est dérivable sur $\mathbb{R}^+$ comme composée de $T \mapsto a\left(\frac{1}{T} - \frac{1}{T_1}\right)$, dérivable sur $\mathbb{R}^+$ car c'est une fraction rationnelle sans pôle sur $\mathbb{R}^+$, par $x \mapsto e^x$, dérivable sur $\mathbb{R}$. Dérivons $R(T)$:

$$\frac{dR}{dT}(T) = -a \frac{1}{T^2} R_1 e^{\left[a\left(\frac{1}{T} - \frac{1}{T_1}\right)\right]} = -a \frac{1}{T^2} R(T).$$

Par conséquent, $k(T) = -\frac{a}{T^2}$ et à $T_1 = 300 \text{ K}$, nous avons $k(T_1) \approx -4,44 \cdot 10^{-2} \text{ K}^{-1}$.

2. Si nous mesurons la résistance R à 10^{-3} près, nous pouvons détecter ΔT :

$$\Delta T = \left| \frac{1}{k} \right| \frac{\Delta R}{R} \approx 2,3 \cdot 10^{-2} \text{ K}.$$

Pour s'entraîner (solutions p. 523)

37.1 Calculer la dérivée de la fonction f définie sur $\mathbb{R}$ par : $f(x) = \frac{x^2 + 3}{2x^4 + 8}$.

37.2 Calculer la dérivée troisième de la fonction f définie sur $\mathbb{R}$ par : $f(x) = \cos(7x + 2)e^{-x^2}$.

37.3 Dériver la fonction f définie sur $\mathbb{R}$ par : $f(x) = \frac{x^3 + 4x^2 - 11}{x^4 + x^2 + 1}$.

37.4 Dériver la fonction f définie sur $]2; +\infty[$ par $f(x) = 3 \frac{\ln(x^2 - 4)}{\sqrt{x^3 + 4}}$.

37.5 Déterminer, pour tout $n \in \mathbb{N}$, la dérivée $n^{\text{ème}}$ de la fonction f définie sur $\mathbb{R}$ par : $f(x) = (x^3 + x^2 + 1) \exp(-x)$.

37.6 Déterminer, pour tout $n \in \mathbb{N}$, la dérivée $n^{\text{ème}}$ de la fonction f définie sur $] -1; +1[$ par : $f(x) = \frac{2x + 3}{(x - 1)^2(x + 1)}$.

37.7 

Soit f une fonction définie sur $] -1; +1[$ à valeurs dans $\mathbb{R}$ définie par : $f(x) = \frac{1}{\sqrt{1 - x^2}}$.

1. Montrer que f est indéfiniment dérivable sur $] -1; +1[$ et que, pour tout $n \in \mathbb{N}$, il existe un polynôme P_n de $\mathbb{R}[X]$ tel que :

$$\forall x \in] -1; +1[, \quad f^{(n)}(x) = \frac{P_n(x)}{(1 - x^2)^{n+\frac{1}{2}}}.$$

Établir :

$$\forall n \in \mathbb{N}, P_{n+1} = (1 - X^2)P'_n + (2n + 1)XP_n.$$

2.a. Montrer :

$$\forall x \in] -1; +1[, \quad (1 - x^2)f'(x) - xf(x) = 0.$$

2.b. En déduire, pour tout $n \in \mathbb{N}^*$:

$$P_{n+1} - (2n + 1)XP_n - n^2(1 - X^2)P_{n-1} = 0.$$

2.c. Montrer : $\forall n \in \mathbb{N}^*, P'_n = n^2P_{n-1}$.

2.d. En déduire, pour tout $n \in \mathbb{N} \setminus \{0; 1\}$:

$$n^2P_n - (2n - 1)XP'_n - (1 - X^2)P''_n = 0.$$

3. Calculer $P_n(0)$ pour tout $n \in \mathbb{N}$.

Étude globale des fonctions dérivables

1. Théorème de Rolle et des accroissements finis

Théorème 38.1 Théorème de Rolle

Soit $f : [a; b] \rightarrow \mathbb{R}$ une fonction continue sur un intervalle $[a; b]$ avec $a < b$, dérivable sur $]a; b[$ et telle que $f(a) = f(b)$. Alors il existe au moins un point $c \in]a; b[$ tel que $f'(c) = 0$.

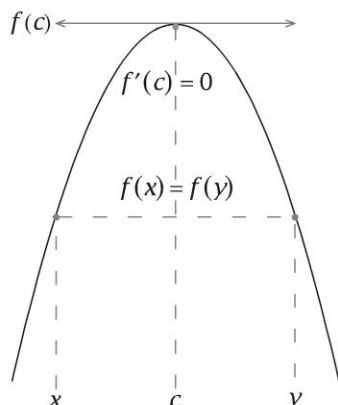


Illustration du théorème de Rolle

Théorème 38.2 Théorème des accroissements finis

Soit $f : [a; b] \rightarrow \mathbb{R}$ une fonction continue sur un intervalle $[a; b]$ avec $a < b$, dérivable sur $]a; b[$. Alors il existe au moins un point $c \in]a; b[$ tel que $f(b) - f(a) = f'(c) \times (b - a)$.

Théorème 38.3 Inégalité des accroissements finis

Soit $f : [a; b] \rightarrow \mathbb{R}$ une fonction continue sur un intervalle $[a; b]$ avec $a < b$, dérivable sur $]a; b[$. Si m et M sont deux réels tels que pour tout $x \in]a; b[$, $m \leq f'(x) \leq M$ alors nous avons : $m \times (b - a) \leq f(b) - f(a) \leq M \times (b - a)$. En particulier si M est un réel tel que pour tout $x \in]a; b[$, $|f'(x)| \leq M$ alors nous avons $|f(b) - f(a)| \leq M \times |b - a|$.

Théorème 38.4 Théorème « limite de la dérivée »

Soient I un intervalle de $\mathbb{R}$ tel que $x_0 \in I$ et $f : I \rightarrow \mathbb{R}$. Si f est continue en x_0 , dérivable sur $I \setminus \{x_0\}$ et si f' admet une limite finie l en x_0 , alors f est dérivable en x_0 .

2. Fonctions dérivables et fonctions monotones

Définition 38.1

Soient D une partie de $\mathbb{R}$ et $f : D \rightarrow \mathbb{R}$ une fonction.

- La fonction f est **croissante** (strictement croissante) sur D si :

$$\forall (x, y) \in D^2, \quad x \leq y \implies f(x) \leq (<) f(y).$$

- La fonction f est **décroissante (strictement décroissante)** sur D si :

$$\forall (x, y) \in D^2, \quad x \leq y \implies f(x) \geq (>) f(y).$$

- La fonction f est **monotone** sur D si elle est croissante sur D ou décroissante sur D .
- La fonction f est **strictement monotone** sur D si elle est strictement croissante sur D ou strictement décroissante sur D .

Théorème 38.5

Soit une fonction $f : I \longrightarrow \mathbb{R}$ dérivable sur un intervalle ouvert I .

- Si pour tout $x \in I$, $f'(x) \geq 0$, alors la fonction f est croissante sur l'intervalle I .
- Si pour tout $x \in I$, $f'(x) \leq 0$, alors la fonction f est décroissante sur l'intervalle I .
- Si pour tout $x \in I$, $f'(x) = 0$, alors la fonction f est constante sur l'intervalle I .
- Si pour tout $x \in I$, $f'(x) > 0$, alors la fonction f est strictement croissante sur l'intervalle I .
- Si pour tout $x \in I$, $f'(x) < 0$, alors la fonction f est strictement décroissante sur l'intervalle I .

3. Extrema pour une fonction dérivable à valeurs réelles

Définition 38.2

- Un point x_0 de $]a; b[$ est un **extremum local** si et seulement si $f'(x_0) = 0$ et f' change de signe en x_0 .
- Un point x_0 de $]a; b[$ est un **minimum local** si de plus $f''(x_0) \geq 0$.
- Un point x_0 de $]a; b[$ est un **maximum local** si de plus $f''(x_0) \leq 0$.

4. Convexité

Définition 38.3

Soit I un intervalle.

- Une **fonction** $f : I \longrightarrow \mathbb{R}$ est **convexe** sur l'intervalle I si la partie du plan située au-dessus de la courbe est convexe ; c'est-à-dire si tout arc de sa courbe représentative est situé au-dessous de la corde correspondante.

Ceci se traduit par :

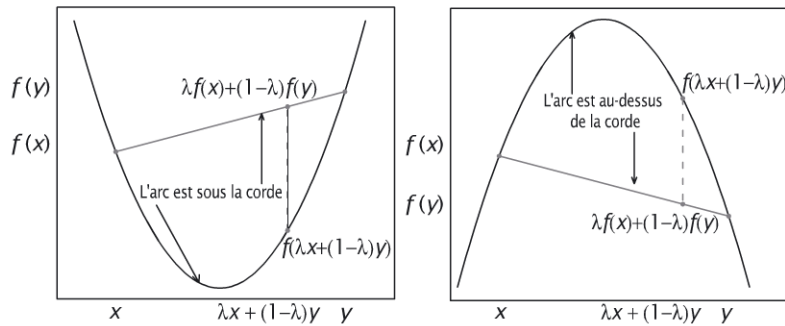
$$\forall x_1 \in I, \forall x_2 \in I, \forall \alpha \in [0; 1], f(\alpha x_1 + (1 - \alpha)x_2) \leq \alpha f(x_1) + (1 - \alpha)f(x_2).$$

- Une **fonction** $f : I \longrightarrow \mathbb{R}$ est **concave** sur l'intervalle I si la fonction $-f$ est convexe.

Théorème 38.6 Inégalité de convexité

Soit $f : I \longrightarrow \mathbb{R}$ une fonction convexe sur un intervalle I . Si les points $x_1, \dots, x_n$ appartiennent à l'intervalle I , si $\alpha_1, \dots, \alpha_n$ sont des réels positifs tels que $\sum_{i=1}^n \alpha_i = 1$, alors nous avons :

$$f\left(\sum_{i=1}^n \alpha_i x_i\right) \leq \sum_{i=1}^n \alpha_i f(x_i).$$



Fonctions convexe et concave

Proposition 38.1 Lemme des trois cordes

Si la fonction f est convexe sur l'intervalle I , alors nous avons, pour tous points x_1, x_2, x_3 de I avec $x_1 < x_2 < x_3$:

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq \frac{f(x_3) - f(x_1)}{x_3 - x_1} \leq \frac{f(x_3) - f(x_2)}{x_3 - x_2}.$$

Théorème 38.7

Si la fonction f est convexe sur un intervalle ouvert I , alors

- la fonction f est continue en tout point de l'intervalle I ;
- la fonction f est dérivable à gauche et à droite en tout point de l'intervalle I , et les fonctions f'_g et f'_d sont croissantes sur l'intervalle I ;
- l'ensemble des points x où la fonction f n'est pas dérivable (c'est-à-dire tels que $f'_g(x) \neq f'_d(x)$) est au plus dénombrable.

Théorème 38.8

Soit une fonction $f : I \rightarrow \mathbb{R}$ dérivable sur un intervalle I . La fonction f est convexe si et seulement si la fonction f' est croissante.

Corollaire 38.1

Si la fonction $f : I \rightarrow \mathbb{R}$ est deux fois dérivable sur l'intervalle I , alors la fonction f est convexe si et seulement si la fonction f'' est positive sur l'intervalle I .

Application

1. Montrer qu'entre deux solutions réelles de $e^x \sin(x) = 1$, il existe au moins une solution réelle de $e^x \cos(x) = -1$.

2. Montrer pour tout $x > 0$ que $\frac{1}{2\sqrt{x+1}} \leq \sqrt{x+1} - \sqrt{x} \leq \frac{1}{2\sqrt{x}}$.

3. Montrer que $x_0 = -7/2$ est un minimum local de la fonction f définie sur $\mathbb{R}$ par $f(x) = x^2 + 7x + 3$.

4. Hideki Yukawa a démontré que l'atome d'hydrogène, de rayon noté a_0 , pouvait être représenté par un proton immobile et un électron mobile dont la probabilité de présence sur une sphère de rayon de r autour de l'atome pouvait être écrite $\rho(r) = \frac{qr}{a_0^2} e^{-\frac{r}{a_0}}$. Pour quelle valeur de r , la probabilité de présence $\rho(r)$ est-elle maximum ?

1. Soient a, b deux solutions réelles de l'équation $e^x \sin(x) = 1$ avec $a < b$. Introduisons la fonction f définie sur $\mathbb{R}$ par : $f(x) = e^{-x} - \sin(x)$. f est continue sur $\mathbb{R}$ donc sur $[a; b]$, dérivable sur $\mathbb{R}$ donc sur $]a; b[$. De plus, nous avons : $f(a) = \frac{1}{e^a} - \sin(a) = 0$ car $e^a \sin(a) = 1$ et $f(b) = \frac{1}{e^b} - \sin(b) = 0$ car $e^b \sin(b) = 1$. Donc $f(a) = f(b) = 0$. Donc d'après le théorème de Rolle, il existe $c \in]a; b[$ tel que $f'(c) = 0$. Or $f'(x) = -e^{-x} - \cos(x)$ donc $f'(c) = 0 \iff -e^{-c} - \cos(c) = 0 \iff e^c \cos(c) = -1$. L'équation $e^x \cos(x) = -1$ admet donc une solution réelle dans $]a; b[$.

2. Soit la fonction f définie sur $\mathbb{R}_+^*$ par $f(t) = \sqrt{t}$. La fonction f est dérivable sur $\mathbb{R}_+^*$ et $\forall t \in \mathbb{R}_+^*, f'(t) = \frac{1}{2\sqrt{t}}$. Soit $x > 0$ tel que $0 < x \leq t \leq x+1$. Alors nous avons : $0 < \sqrt{x} \leq \sqrt{t} \leq \sqrt{x+1}$, ce qui implique : $\frac{1}{2\sqrt{x+1}} \leq \frac{1}{2\sqrt{t}} \leq \frac{1}{2\sqrt{x}}$ pour $x > 0$ et $t \in [x; x+1]$. En utilisant l'inégalité des accroissements finis sur $[x; x+1]$ avec $x > 0$, nous avons : $\frac{x+1-x}{2\sqrt{x+1}} \leq f(x+1) - f(x) \leq \frac{x+1-x}{2\sqrt{x}}$. Ainsi pour tout $x > 0$ nous avons : $\frac{1}{2\sqrt{x+1}} \leq \sqrt{x+1} - \sqrt{x} \leq \frac{1}{2\sqrt{x}}$.

3. La fonction f est dérivable sur $\mathbb{R}$ comme fonction polynomiale. Nous avons : $\forall x \in \mathbb{R}, f'(x) = 2x + 7$, donc $f'(x_0) = 0 \iff x_0 = -7/2$. De plus, pour tout $x \in \mathbb{R}$ nous avons : $f''(x) = 2 > 0$. Donc $x_0 = -7/2$ est bien un minimum local pour f .

4. Elle est maximum pour $r = a_0$, le rayon de l'atome.

Pour s'entraîner (solutions p. 523)

38.1 Soit la fonction f définie sur $\mathbb{R}$ par : $f(x) = x^6 - x^3 + 4$. Montrer qu'il existe un réel $c \in]0; 1[$ tel que $f'(c) = 0$.

38.2 Soient f une fonction définie sur $[0; 1]$ à valeurs réelles convexe et ϕ la fonction définie sur $\left[0; \frac{1}{2}\right]$ à valeurs réelles par $\phi(x) = f(x) + f(1-x)$. Montrer que ϕ est décroissante.

38.3 

1. Montrer : $\forall x \in]-1; +\infty[, \frac{x}{1+x} \leq \ln(1+x) \leq x$.

2. Montrer : $\forall k \in \mathbb{N}^*, \left(1 + \frac{1}{k}\right)^k < e < \left(1 + \frac{1}{k}\right)^{k+1}$.

3. En déduire : $\forall n \in \mathbb{N}^*, \sum_{k=1}^n \frac{\ln(k^2 + k)}{k+1} < (\ln(n+1))^2 < \sum_{k=1}^n \frac{\ln(k^2 + k)}{k}$.

38.4 Soient $(a, b) \in \mathbb{R}^2$, tels que $a < b$, $f, g : [a; b] \rightarrow \mathbb{R}$ continues sur $[a; b]$, dérivables sur $]a; b[$, telles que : $\forall x \in]a; b[, g'(x) \neq 0$. Montrer qu'il existe $c \in]a; b[$ tel que : $\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(c)}{g'(c)}$. Indication : considérer la fonction $h(x) = (f(b) - f(a)) \times (g(x) - g(b)) - (g(b) - g(a)) \times (f(x) - f(b))$.

38.5 Montrer que la fonction f définie sur $\mathbb{R}$ par : $f(x) = (x^2 + 4)e^x$ est convexe sur $\mathbb{R}$.

38.6 

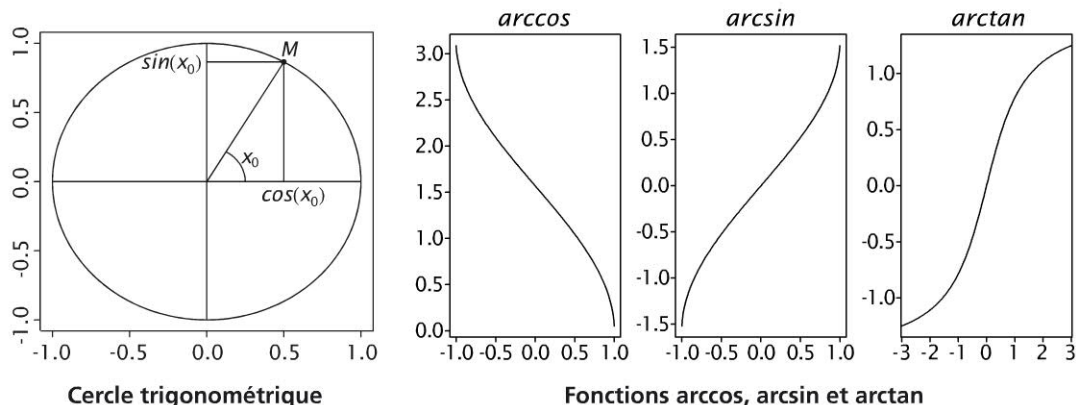
1. Montrer que, au voisinage de 0 : $-\frac{x^2}{2} - 3x^4 \leq \ln(\cos(x)) \leq -\frac{x^2}{2}$.

2. En déduire $\lim_{n \rightarrow +\infty} \prod_{k=1}^n \cos\left(\frac{\sqrt{k}}{n}\right)$.

Fonctions circulaires et circulaires réciproques

Les fonctions circulaires, appelées également fonctions trigonométriques, sont des fonctions dont la variable est un angle.

1. Fonctions cosinus, sinus et tangente



Définition 39.1

Dans le plan muni du repère orthonormé, le **cercle trigonométrique** est le cercle de centre O et de rayon 1. Soient M un point de ce cercle, et x_0 l'angle $(\vec{i}; \widehat{OM}) \in [0; 2\pi]$ (Figure 1). Le **cosinus** et le **sinus** de x_0 , notés respectivement $\cos(x_0) \in [-1; 1]$ et $\sin(x_0) \in [-1; 1]$ sont définis par les coordonnées du point M dans le repère $(O, \vec{i}, \vec{j})$. Cette définition de cosinus et sinus s'étend par 2π -périodicité sur $\mathbb{R}$ tout entier : $\forall x \in \mathbb{R}, \forall k \in \mathbb{Z}, \cos(x) = \cos(x + 2k\pi)$ et $\sin(x) = \sin(x + 2k\pi)$.

Les fonctions cos et sin vérifient les propriétés suivantes :

- Les fonctions cos et sin sont à valeurs dans $[-1; 1]$.
- La fonction cos est paire, et la fonction sin est impaire : $\forall x \in \mathbb{R}, \cos(-x) = \cos x$ et $\sin(-x) = -\sin x$.
- Les fonctions cos et sin sont dérivables sur $\mathbb{R}$: $\forall x \in \mathbb{R}, \cos'(x) = -\sin x$ et $\sin'(x) = \cos x$.
- Pour tout $x \in \mathbb{R}$:

$$\begin{aligned}\cos^2(x) + \sin^2(x) &= 1 \\ \cos(2x) &= \cos^2(x) - \sin^2(x) = 2\cos^2(x) - 1 = 1 - 2\sin^2(x) \\ \sin(2x) &= 2\sin(x)\cos(x).\end{aligned}$$

Définition 39.2

La fonction **tangente** est définie sur $\mathbb{R} \setminus \left\{ \frac{\pi}{2} + k\pi, k \in \mathbb{Z} \right\}$ par $\tan(x) = \frac{\sin(x)}{\cos(x)}$. La fonction tangente est π -périodique, impaire, dérivable sur son domaine de définition et sa dérivée est égale à $\tan'(x) = 1 + \tan^2(x) = \frac{1}{\cos^2(x)}$.

2. Fonctions arcsinus, arccosinus, arctangente**Définition 39.3**

- La fonction cos est continue, strictement décroissante de $[0; \pi]$ sur $[-1; 1]$. Ainsi la fonction cos définit une bijection de $[0; \pi]$ sur $[-1; 1]$. Sa fonction réciproque est appelée **fonction arccosinus** et notée $\arccos(x)$. Nous avons

$$y = \arccos(x) \quad (x \in [-1; 1]) \iff x = \cos(y) \quad (y \in [0; \pi]).$$

La fonction arccos est dérivable sur $] -1; 1[$ et $\forall x \in] -1; 1[, \arccos'(x) = -\frac{1}{\sqrt{1-x^2}}$.

- La fonction sin est continue, strictement croissante de $\left[-\frac{\pi}{2}; \frac{\pi}{2}\right]$ sur $[-1; 1]$. Ainsi la fonction sin définit une bijection de $\left[-\frac{\pi}{2}; \frac{\pi}{2}\right]$ sur $[-1; 1]$. Sa fonction réciproque est appelée **fonction arcsinus** et notée $\arcsin(x)$. Nous avons :

$$y = \arcsin(x) \quad (x \in [-1; 1]) \iff x = \sin(y) \quad \left(y \in \left[-\frac{\pi}{2}; \frac{\pi}{2}\right]\right).$$

La fonction arcsin est dérivable sur $] -1; 1[$ et $\forall x \in] -1; 1[, \arcsin'(x) = \frac{1}{\sqrt{1-x^2}}$.

- La fonction tan est continue, strictement croissante de $\left]-\frac{\pi}{2}; \frac{\pi}{2}\right[$ sur $\mathbb{R}$. Ainsi, la fonction tan définit une bijection de $\left]-\frac{\pi}{2}; \frac{\pi}{2}\right[$ sur $\mathbb{R}$. Sa fonction réciproque est appelée **fonction arctangente** et notée $\arctan(x)$. Nous avons :

$$y = \arctan(x) \quad (x \in \mathbb{R}) \iff x = \tan(y) \quad \left(y \in \left]-\frac{\pi}{2}; \frac{\pi}{2}\right[\right).$$

La fonction arctan est dérivable sur $\mathbb{R}$ et $\forall x \in \mathbb{R}, \arctan'(x) = \frac{1}{1+x^2}$.



Nous avons les égalités fondamentales suivantes :

$$\forall x \in [-1; 1], \quad \arccos(-x) + \arccos(x) = \pi \quad \text{et} \quad \arccos(x) + \arcsin(x) = \frac{\pi}{2}.$$

$$\forall x \in]0; +\infty[, \quad \arctan(x) + \arctan\left(\frac{1}{x}\right) = \frac{\pi}{2} \quad \text{et} \quad \forall x \in]-\infty; 0[, \quad \arctan(x) + \arctan\left(\frac{1}{x}\right) = -\frac{\pi}{2}.$$

1. La Terre possède un champ magnétique intrinsèque, appelé champ géomagnétique, provenant du mouvement des particules chargées dans son noyau fluide. En première approximation, ce champ est celui d'un dipôle placé au centre O de la Terre. Le champ magnétique B à la surface de la Terre est de la forme :

$$B(\lambda) = k \frac{\sqrt{1 + 3 \sin^2(\lambda)}}{\cos^6(\lambda)},$$

où λ est la latitude géomagnétique et k une constante positive. Étudier la fonction $B(\lambda)$ pour $\lambda \in \left[-\frac{\pi}{3}; +\frac{\pi}{3}\right]$.

2. Soit un filtre analogique composé d'une résistance et d'un condensateur. Le déphasage du signal à la sortie du filtre par rapport au signal à l'entrée du filtre est égale à :

$$\varphi(f) = -\arctan(2\pi\tau f),$$

où $\tau > 0$ est une constante de temps du filtre et f la fréquence. Étudier $\varphi(f)$ pour $f \in [0; +\infty[$.

1. La fonction B est dérivable sur $\left[-\frac{\pi}{3}; +\frac{\pi}{3}\right]$. Nous obtenons, pour tout $\lambda \in \left[-\frac{\pi}{3}; +\frac{\pi}{3}\right]$:

$$B'(\lambda) = k \left(\frac{3 \cos(\lambda) \sin(\lambda) (1 + 3 \sin^2(\lambda))^{-\frac{1}{2}} \cos^6(\lambda) + \sqrt{1 + 3 \sin^2(\lambda)} 6 \sin(\lambda) \cos^5(\lambda)}{\cos^{12}(\lambda)} \right).$$

Par conséquent, après simplifications, nous avons, pour tout $\lambda \in \left[-\frac{\pi}{3}; +\frac{\pi}{3}\right]$:

$$B'(\lambda) = 3k \left(\frac{\sin(2\lambda) \cos(\lambda) + 4 \sin(\lambda) + 12 \sin^3(\lambda)}{2 \cos^7(\lambda) \sqrt{1 + 3 \sin^2(\lambda)}} \right).$$

D'autre part, nous avons que pour tout $\lambda \in \left[-\frac{\pi}{3}; +\frac{\pi}{3}\right]$, $2 \cos^7(\lambda) \sqrt{1 + 3 \sin^2(\lambda)} > 0$. Il nous reste donc à étudier le numérateur de la fonction B' . Posons $g(\lambda) = \sin(2\lambda) \cos(\lambda) + 4 \sin(\lambda) + 12 \sin^3(\lambda)$. La fonction g est dérivable sur $\left[-\frac{\pi}{3}; +\frac{\pi}{3}\right]$ et nous avons, pour tout $\lambda \in \left[-\frac{\pi}{3}; +\frac{\pi}{3}\right]$:

$$g'(\lambda) = \cos(\lambda) (36 - 30 \cos^2(\lambda)).$$

Or $\cos(\lambda) > 0$ pour tout $\lambda \in \left[-\frac{\pi}{3}; +\frac{\pi}{3}\right]$. De plus, pour tout $\lambda \in \left[-\frac{\pi}{3}; +\frac{\pi}{3}\right]$, nous avons $0 < \cos(\lambda) \leq 1$. Ainsi pour tout $\lambda \in \left[-\frac{\pi}{3}; +\frac{\pi}{3}\right]$, nous avons $36 - 30 \cos^2(\lambda) > 0$ et par conséquent pour tout $\lambda \in \left[-\frac{\pi}{3}; +\frac{\pi}{3}\right]$, $g'(\lambda) > 0$. Donc g est strictement croissante sur $\left[-\frac{\pi}{3}; +\frac{\pi}{3}\right]$. g croît de $-\frac{27}{4} \sqrt{3}$ à $\frac{27}{4} \sqrt{3}$. Donc il existe un unique $\lambda_0 \in \left[-\frac{\pi}{3}; +\frac{\pi}{3}\right]$ tel que $g(\lambda_0) = 0$. Par conséquent, il existe un unique λ_0 tel que $B'(\lambda_0) = 0$. Nous voyons que $\lambda_0 = 0$ est la solution car $\sin(0) = 0$. Donc B est strictement décroissante sur $\left[-\frac{\pi}{3}; 0\right]$ de $32 \sqrt{13}k$ à k et est strictement croissante sur $\left[0; +\frac{\pi}{3}\right]$ de k à $32 \sqrt{13}k$.

2. La fonction φ est dérivable sur $[0; +\infty[$. Ainsi nous avons pour tout $f \in [0; +\infty[$:

$$\varphi'(f) = \frac{-2\pi\tau}{1 + (2\pi\tau f)^2} < 0.$$

φ est strictement décroissante sur $[0; +\infty[$. Elle prend ses valeurs dans l'intervalle $\left]-\frac{\pi}{2}; 0\right]$.

Pour s'entraîner (solutions p. 524)

39.1 1. Dériver la fonction suivante :
 $f(x) = \cos(2x) \sin(3x) + x^2 \tan(8x^2).$

2. Dériver la fonction suivante : $f(x) = (3x + 4)^5(x^2 - \cos(x) + x)^3.$

3. Dériver la fonction suivante : $f(x) = \arctan\left(\frac{x^3 + \cos(x)}{x^2 + 3x + 1}\right).$

39.2 Calculer : $\arcsin\left(\frac{5}{6}\right) + \arcsin\left(\frac{\sqrt{11}}{6}\right).$

39.3 Résoudre l'équation : $\arccos(x) + \arccos(\sqrt{1-x^2}) = \frac{\pi}{2}.$

39.4 1. Soit $x > 0$. Montrer que :
 $\arctan(x) + \arctan\left(\frac{1}{x}\right) = \frac{\pi}{2}.$

2. Soit $x < 0$. Montrer que : $\arctan(x) + \arctan\left(\frac{1}{x}\right) = -\frac{\pi}{2}.$

39.5 Montrer que :

$$\forall x \in \mathbb{R}, \quad \sin(\arctan(x)) = \frac{x}{\sqrt{1+x^2}}.$$

39.6 1. Soit $x \neq \frac{\pi}{2}(k+1), k \in \mathbb{Z}$. Montrer que :

$$\tan(x) = -2 \cotan(2x) + \cotan(x)$$

$$\text{où } \cotan(x) = \frac{\cos(x)}{\sin(x)}.$$

2. Puis calculer, pour $n \geq 1$:

$$S_n = \tan(x) + 2 \tan(2x) + 2^2 \tan(2^2x) + \dots + 2^n \tan(2^n x).$$

39.7 1. Soit $x \neq \frac{\pi}{2}(k+1), k \in \mathbb{Z}$. Montrer que :

$$1 + \frac{1}{\cos(2x)} = \frac{\tan(2x)}{\tan(x)}.$$

2. Calculer, pour $n \geq 1$:

$$S_n = \left(1 + \frac{1}{\cos(x)}\right) \left(1 + \frac{1}{\cos(2x)}\right) \dots \left(1 + \frac{1}{\cos(2^n x)}\right).$$

39.8 Soit $a = \arcsin\left(\frac{\sqrt{5}+1}{4}\right).$

1. Calculer $\cos(2a)$ et $\cos(4a)$.

2. En déduire a .

39.9 

Montrer les inégalités suivantes :

• $\forall (x, y) \in \mathbb{R}^2$

$$\left(0 < x < y < \frac{\pi}{4} \implies \frac{y}{x} < \frac{\tan(y)}{\tan(x)} < \frac{4y}{\pi x}\right)$$

• $\forall x \in \left]0; \frac{\pi}{2}\right[, \forall y \in]0; 1[,$

$$x - \arcsin(y) \leq \frac{\sqrt{1-y^2} - \cos(x)}{y}.$$

Fonctions hyperboliques et hyperboliques réciproques

1. Fonctions cosinus, sinus et tangente hyperboliques

Définition 40.1

- La fonction **cosinus hyperbolique**, notée ch , est définie sur $\mathbb{R}$, par :

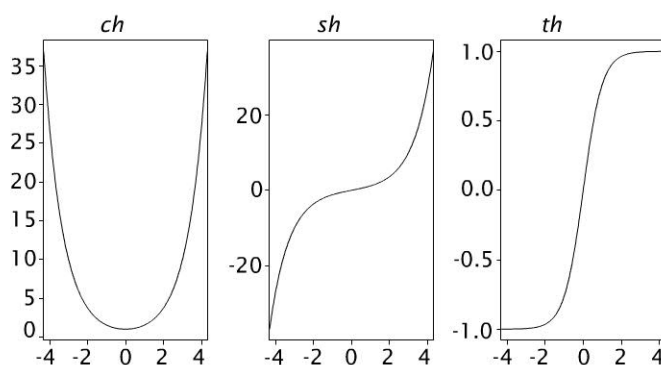
$$\forall x \in \mathbb{R}, \quad \text{ch}(x) = \frac{e^x + e^{-x}}{2}.$$

- La fonction **sinus hyperbolique**, notée sh , est définie sur $\mathbb{R}$, par :

$$\forall x \in \mathbb{R}, \quad \text{sh}(x) = \frac{e^x - e^{-x}}{2}.$$

- La fonction **tangente hyperbolique**, notée th , est définie sur $\mathbb{R}$, par :

$$\forall x \in \mathbb{R}, \quad \text{th}(x) = \frac{\text{sh}(x)}{\text{ch}(x)} = \frac{e^{2x} - 1}{e^{2x} + 1}.$$



Fonctions cosinus, sinus et tangente hyperboliques

Les fonctions ch , sh , th vérifient les propriétés suivantes.

- La fonction ch est paire, et les fonctions sh et th sont impaires : $\forall x \in \mathbb{R}, \text{ch}(-x) = \text{ch } x$, $\text{sh}(-x) = -\text{sh } x$ et $\text{th}(-x) = -\text{th}(x)$.
- Les fonctions ch , sh et th sont dérivables sur $\mathbb{R}$: $\forall x \in \mathbb{R}, \text{ch}'(x) = \text{sh } x$, $\text{sh}'(x) = \text{ch } x$ et $\text{th}'(x) = 1 - \text{th}^2(x) = \frac{1}{\text{ch}^2(x)}$.
- La fonction ch est strictement décroissante sur $\mathbb{R}_-$ et strictement croissante sur $\mathbb{R}_+$. La fonction sh est strictement croissante sur $\mathbb{R}$. La fonction th est strictement croissante sur $\mathbb{R}$.

- Il existe des formules de trigonométrie hyperbolique analogues à celles de la trigonométrie classique. Pour tout $x \in \mathbb{R}$, nous avons :

$$\operatorname{ch}^2(x) - \operatorname{sh}^2(x) = 1$$

$$\operatorname{ch}(2x) = \operatorname{ch}^2(x) + \operatorname{sh}^2(x) = 2 \operatorname{ch}^2(x) - 1 = 1 + 2 \operatorname{sh}^2(x)$$

$$\operatorname{sh}(2x) = 2 \operatorname{sh}(x) \operatorname{ch}(x).$$

2. Fonctions hyperboliques réciproques

Définition 40.2

- La fonction ch est continue, strictement croissante de $[0; +\infty[$ sur $[1; +\infty[$. Ainsi la fonction ch définit une bijection de $[0; +\infty[$ sur $[1; +\infty[$. Sa fonction réciproque est appelée **fonction argument cosinus hyperbolique** et notée $\operatorname{argch}(x)$. Nous avons

$$y = \operatorname{argch}(x) \quad x \in [1; +\infty[\iff x = \operatorname{ch}(y) \quad y \in [0; +\infty[.$$

De plus, nous avons :

$$\forall x \in [1; +\infty[, \quad \operatorname{argch}(x) = \ln \left(x + \sqrt{x^2 - 1} \right).$$

La fonction argch est dérivable sur $]1; +\infty[$ et $\forall x \in]1; +\infty[, \operatorname{argch}'(x) = \frac{1}{\sqrt{x^2 - 1}}$.

La fonction argch est strictement croissante sur $]1; +\infty[$.

- La fonction sh est continue, strictement croissante de $\mathbb{R}$ sur $\mathbb{R}$. Ainsi la fonction sh définit une bijection de $\mathbb{R}$ sur $\mathbb{R}$. Sa fonction réciproque est appelée **fonction argument sinus hyperbolique** et notée $\operatorname{argsh}(x)$. Nous avons

$$y = \operatorname{argsh}(x) \quad (x \in \mathbb{R}) \iff x = \operatorname{sh}(y) \quad (y \in \mathbb{R}).$$

De plus, nous avons :

$$\forall x \in \mathbb{R}, \quad \operatorname{argsh}(x) = \ln \left(x + \sqrt{x^2 + 1} \right).$$

La fonction argsh est dérivable sur $\mathbb{R}$ et $\forall x \in \mathbb{R}, \operatorname{argsh}'(x) = \frac{1}{\sqrt{x^2 + 1}}$. La fonction argsh est strictement croissante sur $\mathbb{R}$.

- La fonction th est continue, strictement croissante de $\mathbb{R}$ sur $] -1; +1[$. Ainsi la fonction th définit une bijection de $] -1; +1[$ sur $\mathbb{R}$. Sa fonction réciproque est appelée **fonction argument tangente hyperbolique** et notée $\operatorname{argth}(x)$. Nous avons

$$y = \operatorname{argth}(x) \quad x \in] -1; +1[\iff x = \operatorname{th}(y) \quad y \in \mathbb{R}.$$

De plus, nous avons :

$$\forall x \in \mathbb{R}, \quad \operatorname{argth}(x) = \frac{1}{2} \ln \left(\frac{1+x}{1-x} \right).$$

La fonction argth est dérivable sur $] -1; +1[$ et $\forall x \in] -1; +1[, \operatorname{argth}'(x) = \frac{1}{1-x^2}$.
La fonction argth est strictement croissante sur $] -1; +1[$.

Application

1. Champ électrique photon masse non nulle

Soit une sphère creuse de rayon R et de centre O , portant une charge uniformément répartie sur sa surface. Le calcul du module du champ électrique $\vec{E}$ à une distance r du centre O est donné par :

$$E(r) = k \left(\frac{\eta \operatorname{ch}(\eta r)}{r} - \frac{\operatorname{sh}(\eta r)}{r^2} \right) \quad \text{pour } r < R \quad \text{et} \quad E(r) = k, \quad \text{pour } r > R$$

où k et η désignent des paramètres strictement positifs, et η est très petit devant 1. Étudier les variations de la fonction E pour $r \in]0; R]$ avec $R = 0,7$.

2. Cylindres conducteurs

Nous avons deux cylindres conducteurs très allongés, identiques, de longueur L et de rayon R , placés parallèlement dans le vide à une distance a . Ces cylindres caractérisent une ligne bifilaire. Nous avons également les inégalités suivantes : $R < a < L$. La capacité linéique du système est égale à :

$$f(L) = \frac{\pi \epsilon_0}{\operatorname{argsh}\left(\frac{L}{2R}\right) - \operatorname{argsh}\left(\frac{L}{2a}\right)}$$

où $\epsilon_0 > 0$ représente la permittivité du vide. Étudier la fonction f .

1. La fonction E est dérivable sur $]0; 0,7[$ et par conséquent nous avons, pour tout $r \in]0; 0,7[$:

$$E'(r) = k \left(\frac{\eta^2 r^2 \operatorname{sh}(\eta r) - 2\eta r \operatorname{ch}(\eta r) + 2 \operatorname{sh}(\eta r)}{r^3} \right).$$



r et k sont strictement positifs, donc la dérivée $E'(r)$ et le numérateur ont le même signe.

Posons $N(r) = \eta^2 r^2 \operatorname{sh}(\eta r) - 2\eta r \operatorname{ch}(\eta r) + 2 \operatorname{sh}(\eta r)$. La fonction N est dérivable sur $]0; 0,7[$ et par conséquent nous avons, pour tout $r \in]0; 0,7[$:

$$N'(r) = 2\eta^2 r \operatorname{sh}(\eta r) + \eta^3 r^2 \operatorname{ch}(\eta r) - 2\eta \operatorname{ch}(\eta r) - 2\eta^2 r \operatorname{sh}(\eta r) + 2\eta \operatorname{ch}(\eta r) = \eta^3 r^2 \operatorname{ch}(\eta r) > 0.$$

Donc la fonction N est strictement croissante sur $]0; 0,7[$. De plus $N(0) = 0$. Ainsi pour tout $r \in]0; 0,7[$, $N(r) \geq 0$ et par conséquent $E'(r) \geq 0$: la fonction E est strictement croissante sur $]0; 0,7[$.



Remarquons que pour calculer la limite de E lorsque r tend vers 0, nous avons besoin d'un développement limité à l'ordre 3 au voisinage de 0 de la fonction sh : $\operatorname{sh}(r) = r + \frac{r^3}{6} + o(r^3)$ et qu'un développement limité à l'ordre 3 au voisinage de 0 de la fonction ch est égal à $\operatorname{ch}(r) = 1 + \frac{r^2}{2} + o(r^3)$.

Ainsi, nous avons :

$$E(r) = k \left(\frac{\eta \left(1 + \frac{(\eta r)^2}{2} + o(r^3) \right)}{r} - \frac{\eta r + \frac{(\eta r)^3}{6} + o(r^3)}{r^2} \right) = k \left(\frac{\eta^3 r}{3} + o(r) \right).$$

Puis $\lim_{r \rightarrow 0} E(r) = 0$. Enfin, $E(0,7) = k \left(\frac{\eta \operatorname{ch}(0,7\eta)}{0,7} - \frac{\operatorname{sh}(0,7\eta)}{0,7^2} \right)$.

2. La fonction argsh étant strictement croissante, et $R < a$, nous avons $\operatorname{argsh}\left(\frac{L}{2R}\right) - \operatorname{argsh}\left(\frac{L}{2a}\right) > 0$, donc f est à valeurs strictement positives.



Fiche 42

La fonction f est dérivable sur $]0; +\infty[$ et pour tout $L > 0$, nous avons :

$$f'(L) = \frac{-\pi\epsilon_0 \left(\frac{1}{2R} \times \frac{1}{\sqrt{1+\frac{L^2}{4R^2}}} - \frac{1}{2a} \times \frac{1}{\sqrt{1+\frac{L^2}{4a^2}}} \right)}{\left(\operatorname{argsh}\left(\frac{L}{2R}\right) - \operatorname{argsh}\left(\frac{L}{2a}\right) \right)^2}.$$



Reste à étudier le signe de $f'(L)$.

$$\text{Nous avons } f'(L) < 0 \iff \frac{1}{R\sqrt{1+\frac{L^2}{4R^2}}} - \frac{1}{a\sqrt{1+\frac{L^2}{4a^2}}} > 0 \iff \frac{1}{\sqrt{R^2+\frac{L^2}{4}}} - \frac{1}{\sqrt{a^2+\frac{L^2}{4}}} > 0.$$

Or, d'après l'énoncé, nous avons $R^2 < a^2$ car $R < a$. Donc f est strictement décroissante sur $]0; +\infty[$.

Maintenant calculons la limite de f quand L tend vers 0. La fonction argsh est continue en 0, donc $\lim_{L \rightarrow 0} \left(\operatorname{argsh}\left(\frac{L}{2R}\right) - \operatorname{argsh}\left(\frac{L}{2a}\right) \right) = 0$ et $\lim_{L \rightarrow 0} f(L) = +\infty$.

Calculons la limite de f quand L tend vers $+\infty$. Nous avons $\operatorname{argsh}\left(\frac{L}{2R}\right) -$

$$\operatorname{argsh}\left(\frac{L}{2a}\right) = \ln\left(\frac{L}{2R} + \sqrt{\frac{L^2}{4R^2} + 1}\right) - \ln\left(\frac{L}{2a} + \sqrt{\frac{L^2}{4a^2} + 1}\right) = \ln\left(\frac{\frac{L}{2R} + \sqrt{\frac{L^2}{4R^2} + 1}}{\frac{L}{2a} + \sqrt{\frac{L^2}{4a^2} + 1}}\right) =$$

$$\ln\left(\frac{\frac{L}{2R} \left(1 + \sqrt{1 + \frac{4R^2}{L^2}}\right)}{\frac{L}{2a} \left(1 + \sqrt{1 + \frac{4a^2}{L^2}}\right)}\right) = \ln\left(\frac{a}{R} \frac{1 + \sqrt{1 + \frac{4R^2}{L^2}}}{1 + \sqrt{1 + \frac{4a^2}{L^2}}}\right) \rightarrow \ln\left(\frac{a}{R}\right) \text{ si } L \rightarrow +\infty. \text{ Par conséquent,}$$

$$\lim_{L \rightarrow +\infty} f(L) = \frac{\pi\epsilon_0}{\ln\left(\frac{a}{R}\right)}.$$

x	0	$+\infty$
$f'(x)$		-
f	$+\infty$	$\frac{\pi\epsilon_0}{\ln\left(\frac{a}{R}\right)}$

Pour s'entraîner (solutions p. 525)

40.1 Déterminer l'ensemble des nombres réels solutions de l'équation :

$$\operatorname{ch}^4(x) + \operatorname{ch}^3(x) \operatorname{sh}(x) + \operatorname{ch}(x) \operatorname{sh}^3(x) + \operatorname{sh}^4(x) = 0.$$

40.2 Déterminer l'ensemble des nombres réels solutions de l'équation :

$$\arctan(\operatorname{ch}(x)) = \arctan\left(\frac{1}{\operatorname{sh}(x)}\right).$$

40.3 Déterminer l'ensemble des nombres réels solutions de l'équation :

$$2 \arctan(\operatorname{th}(x)) = \arctan(\operatorname{sh}(2x)).$$

Formules de Taylor

Dans cette Fiche, I désigne un intervalle de $\mathbb{R}$ ouvert et n un entier naturel non nul.

1. Énoncé de Taylor

Soit $f : I \rightarrow \mathbb{R}$ une fonction dérivable jusqu'à l'ordre n au point $x_0 \in I$. Pour tout $x \in I$, l'expression :

$$f(x) = f(x_0) + \frac{f'(x_0)}{1!}(x - x_0) + \frac{f^{(2)}(x_0)}{2!}(x - x_0)^2 + \cdots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + R_n(x)$$

définit un **reste** $R_n(x)$ dont le comportement s'apparente au monôme $(x - x_0)^{n+1}$.



En présentant cette formule, Taylor propose une méthode de développement en série, mais il se préoccupe peu de la nature du reste $R_n(x)$ qui sera caractérisé ultérieurement, par exemple avec la formule de Taylor avec reste intégral. Nous donnons ci-dessous des expressions caractérisant le reste $R_n(x)$.



Fiche 51



Fiche 45

► Formule de Taylor-Young

Définition 41.1

La fonction ε définie sur $I \setminus \{x_0\}$ par la relation $\varepsilon(x) = \frac{R_n(x)}{(x - x_0)^n}$ vérifie $\lim_{x \rightarrow x_0} \varepsilon(x) = 0$.

La quantité $(x - x_0)^n \varepsilon(x)$ s'appelle le **reste de Young**.



- Le reste de Young se note également $o((x - x_0)^n)$. Nous avons donc $\lim_{x \rightarrow x_0} \frac{o((x - x_0)^n)}{(x - x_0)^n} = 0$.
- Il peut être commode de prolonger la fonction ε par continuité en x_0 en posant $\varepsilon(x_0) = 0$.

Théorème 41.1 Formule de Taylor-Young

Soit $f : I \rightarrow \mathbb{R}$ une fonction définie et dérivable n fois sur un intervalle ouvert I contenant x_0 . Alors, pour $x \in I$, nous avons :

$$f(x) = f(x_0) + \frac{f'(x_0)}{1!}(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \cdots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + o((x - x_0)^n).$$

où $\lim_{x \rightarrow x_0} \frac{o((x - x_0)^n)}{(x - x_0)^n} = 0$. Cette expression s'appelle le **développement limité de f d'ordre n en x_0** .



Cette formule va être très souvent utilisée en $x_0 = 0$.

► Formule de Taylor-Lagrange

Théorème 41.2

Soient $x_0 \in I$ et $f : I \rightarrow \mathbb{R}$ une fonction dérivable sur I jusqu'à l'ordre $n + 1$, alors pour tout $x \in I$, il existe un nombre c strictement compris entre x_0 et x tel que le reste de Young s'écrit :

$$R_n(x) = \frac{f^{(n+1)}(c)}{(n+1)!}(x - x_0)^{n+1}.$$



La formule de Taylor-Lagrange est une généralisation du théorème des accroissements finis.

Fiche 38

Théorème 41.3 Inégalité de Taylor-Lagrange

Soit $f : I \rightarrow \mathbb{R}$ une fonction de classe $\mathcal{C}^{n+1}$ sur un intervalle I de $\mathbb{R}$. S'il existe un réel M tel que $|f^{(n+1)}(y)| \leq M$ pour tout $y \in I$, alors :

$$|R_n(x)| \leq M \frac{|x - x_0|^{n+1}}{(n+1)!}.$$



L'inégalité de Taylor-Lagrange est une conséquence de la formule de Taylor-Lagrange.

2. Obtention de développements limités

Fiche 42

Un grand nombre de fonctions usuelles sont indéfiniment et continûment dérivables sur un intervalle I admettant x_0 comme point intérieur. Nous pouvons alors appliquer la **formule de Taylor-Young** à une telle fonction f , à un ordre quelconque, et en particulier au point $x_0 = 0$.

e^x	$1 + \frac{x}{1!} + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + o(x^n)$
e^{-x}	$1 - \frac{x}{1!} + \frac{x^2}{2!} + \cdots + \frac{(-1)^n x^n}{n!} + o(x^n)$

La deuxième ligne est obtenue en changeant x en $-x$. Puis, $\forall a > 0$, nous avons $a^x = e^{x \ln a}$. Pour α entier positif, le développement de $(1 + e)^a$ coïncide avec la formule du binôme de Newton. Donc :

a^x	$1 + \frac{\ln a}{1!} x + \frac{(\ln a)^2}{2!} x^2 + \cdots + \frac{(\ln a)^n}{n!} x^n + o(x^n)$
$(1+x)^\alpha$	$1 + \frac{\alpha}{1!} x + \frac{\alpha(\alpha-1)}{2!} x^2 + \cdots + \frac{\alpha(\alpha-1) \cdots (\alpha-n+1)}{n!} x^n + o(x^n)$
$\frac{1}{1-x}$	$1 + x + x^2 + \cdots + x^n + o(x^n)$
$\frac{1}{1+x}$	$1 - x + x^2 + \cdots + (-1)^n x^n + o(x^n)$
$\ln(1+x)$	$x - \frac{x^2}{2} + \frac{x^3}{3} + \cdots + \frac{(-1)^{n-1}}{n} x^n + o(x^n)$
$\sin x$	$\frac{x}{1!} - \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots + \frac{(-1)^p}{(2p+1)!} x^{2p+1} + o(x^{2p+2})$

cos x	$1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots + \frac{(-1)^p}{(2p)!} x^{2p} + o(x^{2p+1})$
tan x	$x + \frac{x^3}{3} + \frac{2}{15}x^5 + o(x^6)$
arcsin x	$x + \frac{x^3}{6} + \frac{3}{40}x^5 + o(x^6)$
arccos x	$\frac{\pi}{2} - x - \frac{x^3}{6} - \frac{3}{40}x^5 + o(x^6)$
arctan x	$x - \frac{x^3}{3} + \frac{x^5}{5} - \dots + \frac{(-1)^p}{2p+1} x^{2p+1} + o(x^{2p+2})$
sh x	$\frac{x}{1!} + \frac{x^3}{3!} + \frac{x^5}{5!} + \dots + \frac{x^{2p+1}}{(2p+1)!} + o(x^{2p+2})$
ch x	$1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \dots + \frac{x^{2p}}{(2p)!} + o(x^{2p+1})$
th x	$x - \frac{x^3}{3} + \frac{2}{15}x^5 + o(x^6)$
argsh x	$x - \frac{x^3}{6} + \frac{3}{40}x^5 + o(x^6)$
argth x	$x + \frac{x^3}{3} + \frac{x^5}{5} + \dots + \frac{x^{2p+1}}{2p+1} + o(x^{2p+2})$

Application

En théorie de la relativité restreinte d'Einstein, la masse d'un objet qui se déplace à la vitesse v est : $m = \frac{m_0}{\sqrt{1 - \frac{v^2}{c^2}}}$ où m_0 est la masse de l'objet au repos et c la vitesse de la lumière.

L'énergie cinétique de l'objet est la différence entre son énergie totale et son énergie au repos : $E_c = mc^2 - m_0c^2$.

1. Montrer que si v est petit devant c , c'est-à-dire quand v/c tend vers 0, $E_c = E_{c_N} + o\left(\frac{v^2}{c^2}\right)$ où $E_{c_N} = \frac{1}{2}m_0v^2$, formule de la physique newtonienne classique.

2. Estimer la différence entre les deux expressions E_c et E_{c_N} quand $|v| \leq 1000 \text{ m s}^{-1}$.

1.



Commençons par écrire un développement limité de la fonction E_c à l'ordre 3 en $\frac{v^2}{c^2}$ voisin de 0.

$$\text{Nous avons : } E_c = \frac{m_0c^2}{\sqrt{1 - \frac{v^2}{c^2}}} - m_0c^2 = m_0c^2 \left(\left(1 - \frac{v^2}{c^2}\right)^{-\frac{1}{2}} - 1 \right).$$

$$\text{Comme } \frac{v}{c} \text{ est proche de zéro, } \left(1 - \frac{v^2}{c^2}\right)^{-\frac{1}{2}} = 1 + \frac{v^2}{2c^2} + \frac{\left(-\frac{1}{2}\right)\left(-\frac{3}{2}\right)v^4}{2!c^4} + \frac{\left(-\frac{1}{2}\right)\left(-\frac{3}{2}\right)\left(-\frac{5}{2}\right)\left(-\frac{v^2}{c^2}\right)^3}{3!} + o\left(\left(\frac{v^2}{c^2}\right)^3\right)$$

Donc, $E_c = m_0 c^2 \left(\frac{v^2}{2c^2} + \frac{3v^4}{8c^4} + \frac{5v^6}{16c^6} + o\left(\frac{v^6}{c^6}\right) \right)$.

Or, à des vitesses très petites devant celle de la lumière, comme en mécanique classique, les termes $\frac{v^4}{c^4}$ et $\frac{v^6}{c^6}$ sont négligeables devant $\frac{v^2}{c^2}$. Nous pouvons nous contenter d'un développement limité à l'ordre 1 en $\frac{v^2}{c^2}$. Donc, la partie principale du développement limité à l'ordre 1 de E_c en $\frac{v^2}{c^2}$, qui est égale à $\frac{1}{2}m_0 v^2$, fournit une très bonne approximation pour E_c : $E_c = E_{c_N} + o\left(\frac{v^2}{c^2}\right)$.

2.



Il s'agit de contrôler le reste $E_c - E_{c_N}$ dans le développement limité précédent.

Nous voulons appliquer l'inégalité de Taylor-Lagrange et de ce fait recherchons M un majorant de $|f'''(x)|$ avec $f(x) = m_0 c^2 \left((1+x)^{-\frac{1}{2}} - 1 \right)$ où nous avons posé $x = -\frac{v^2}{c^2}$.

Or, $f'(x) = -\frac{1}{2}m_0 c^2 (1+x)^{-\frac{3}{2}}$ et $f''(x) = \frac{3}{4}m_0 c^2 (1+x)^{-\frac{5}{2}}$, donc comme $v < 1000 \text{ m s}^{-1}$, nous avons : $|f''(x)| \leq \frac{3m_0 c^2}{4 \left(1 - \frac{1000^2}{c^2}\right)^{\frac{5}{2}}} = M$.

En utilisant l'inégalité de Taylor-Lagrange en $x = -\frac{v^2}{c^2}$ et $n = 1$, le reste $E_c - E_{c_N}$ est majoré par $M \frac{|x|^{n+1}}{(n+1)!} = \frac{M}{2} \left| \frac{v^2}{c^2} \right|^2 = 4,16 \cdot 10^{-6} m_0$ avec $v = 1000 \text{ m s}^{-1}$, $c = 3 \cdot 10^8 \text{ m s}^{-1}$.

Donc en approchant E_c par $E_{c_N} = \frac{1}{2}m_0 v^2$, l'erreur est inférieure à $4,16 \times 10^{-6} m_0$ si $v < 1000 \text{ m s}^{-1}$.

Pour s'entraîner (solutions p. 525)

41.1 Montrer que les fonctions définies ci-dessous admettent un développement limité au voisinage de 0 et le calculer.

1. $f(x) = (1-x)^{\frac{1}{2}}$ pour tout n ;
2. $f(x) = \ln(2+x)$ pour tout n ;
3. $f(x) = \sin(x) \cos(2x)$ pour $n = 6$;
4. $f(x) = e^x \ln(1+x)$ pour $n = 3$.

41.2 Soit $a > 0$.

1. Démontrer que :

$$\left| \cos(a) - 1 + \frac{a^2}{2!} - \frac{a^4}{4!} \right| \leq \frac{a^5}{5!}.$$

2. En déduire que :

$$\frac{337}{384} - \frac{1}{3840} \leq \cos(1/2) \leq \frac{337}{384} + \frac{1}{3840}.$$

41.3 Soit x un réel positif.

1. Montrer que : $\left| \ln(1+x) - x + \frac{x^2}{2} \right| \leq \frac{x^3}{3}$.

2. En déduire une valeur approchée de $\ln(1,003)$ à 10^{-8} près.

41.4 Montrer que la suite (u_n) définie par :

$$u_n = 1 - \frac{1}{2} + \frac{1}{3} - \dots + \frac{(-1)^{n-1}}{n} \text{ converge vers } \ln(2).$$

41.5 Soient $f : \mathbb{R} \rightarrow \mathbb{R}$ de classe $\mathcal{C}^2$ et $a \in \mathbb{R}$. Déterminer

$$\lim_{h \rightarrow 0} \frac{f(a+h) + f(a-h) - 2f(a)}{h^2}.$$

Dans cette Fiche, I désigne un intervalle ouvert de $\mathbb{R}$.

Définition 42.1

Soient une fonction $f : I \longrightarrow \mathbb{R}$ et $x_0 \in I$. La fonction f admet un **développement limité d'ordre n au voisinage de x_0** si pour x voisin de x_0 , $f(x)$ s'écrit :

$$f(x) = P_n(x) + (x - x_0)^n \varepsilon(x) \quad \text{avec} \quad \lim_{x \rightarrow x_0} \varepsilon(x) = 0,$$

ou encore :

$$f(x) = P_n(x) + o(x^n) \quad \text{avec} \quad \lim_{x \rightarrow 0} \frac{o(x^n)}{x^n} = 0,$$

et $P_n(x) = \sum_{k=1}^n a_k (x - x_0)^k$ où $a_1, \dots, a_n \in \mathbb{R}$ est la **partie régulière** et $(x - x_0)^n \varepsilon(x)$ est le **reste**.



- Une fonction f définie sur un intervalle I n'admet pas nécessairement un développement limité d'ordre n au voisinage d'un point x_0 de I .
- Si la fonction f admet un développement limité d'ordre n au voisinage de x_0 , ce développement est unique.
- Si la fonction f admet un développement limité d'ordre n au voisinage de x_0 , elle admet un développement limité à tout ordre p inférieur à n au voisinage de x_0 .
- Si la fonction f est dérivable jusqu'à l'ordre n en x_0 , nous obtenons son développement limité d'ordre n au voisinage de x_0 , à partir de la formule de Taylor-Young.

Fiche 41

1. Opérations sur les développements limités

Soient deux fonctions $f, g : I \longrightarrow \mathbb{R}$, avec $0 \in I$, admettant toutes deux un développement limité d'ordre n au voisinage de 0 :

$$\begin{aligned} f(x) &= a_0 + a_1 x + \dots + a_n x^n + o_1(x^n) = P_n(x) + o_1(x^n) \\ g(x) &= b_0 + b_1 x + \dots + b_n x^n + o_2(x^n) = Q_n(x) + o_2(x^n). \end{aligned}$$



Nous avons intentionnellement noté $o_1(x^n)$ et $o_2(x^n)$ les restes des développements limités de f et g , afin d'insister sur le fait que ces restes ne sont pas en général égaux. Ils vérifient tous deux $\lim_{x \rightarrow 0} \frac{o_1(x^n)}{x^n} = \frac{o_2(x^n)}{x^n} = 0$.

► Somme, produit, quotient de développements limités

1. Le développement limité d'ordre n de la **fonction $f + g$** au voisinage de 0 est :

$$f(x) + g(x) = (a_0 + b_0) + (a_1 + b_1)x + \dots + (a_n + b_n)x^n + o_3(x^n) \quad \text{où} \quad \lim_{x \rightarrow 0} \frac{o_3(x^n)}{x^n} = 0$$

2. Considérons le **produit** $f \times g$, nous avons $f(x) \times g(x) = P_n(x) \times Q_n(x) + o_3(x^n)$. Le polynôme $P_n \times Q_n$ est de degré au plus égal à $2n$. Les termes de degré au plus égal à n forment le polynôme :

$$\Pi_n(x) = a_0b_0 + (a_0b_1 + a_1b_0)x + \dots + (a_0b_n + a_1b_{n-1} + \dots + a_nb_0)x^n.$$

Tous les autres termes du produit $P_n \times Q_n$ sont factorisables par x^{n+1} , et finalement nous avons :

$$f(x) \times g(x) = \Pi_n(x) + o_4(x^n) \text{ où } \lim_{x \rightarrow 0} \frac{o_4(x^n)}{x^n} = 0.$$

Ceci est le développement limité d'ordre n du produit $f \times g$ au voisinage de 0.

3. Considérons le **quotient** $\frac{f}{g}$ et supposons que $g(0) = Q_n(0) = b_0 \neq 0$. Nous effectuons la division de $P_n(x)$ par $Q_n(x)$ suivant les puissances croissantes à l'ordre n . Nous obtenons :

$$P_n(x) = Q_n(x) \times T_n(x) + x^{n+1}R(x)$$

où T_n est un polynôme de degré inférieur ou égal à n , et où $R(x)$ est un polynôme. Cette égalité s'écrit : $f(x) - o_1(x^n) = (g(x) - o_2(x^n)) \times T_n(x) + x^{n+1}R(x)$. Nous en déduisons : $f(x) = g(x) \times T_n(x) + o_5(x^n)$. Nous obtenons le développement limité d'ordre n du quotient $\frac{f}{g}$ au voisinage de 0 :

$$\frac{f(x)}{g(x)} = T_n(x) + o_6(x^n) \text{ où } \lim_{x \rightarrow 0} \frac{o_6(x^n)}{x^n} = 0.$$

► Développement limité d'une fonction composée

Si la fonction $g \circ f$ est définie au voisinage de 0 et si $f(0) = 0$, alors la fonction $g \circ f$ admet un développement limité d'ordre n au voisinage de 0. Nous commençons par écrire :

$$g \circ f(x) = Q_n(P_n(x)) + o_2(x^n) = b_0 + b_1P_n(x) + \dots + b_n(P_n(x))^n + o_2(x^n),$$

puis nous ne gardons dans cette expression que les monômes de degré inférieur ou égal à n . Tous les autres, négligeables devant x^n lorsque x est voisin de 0, sont « cachés » dans le reste $o(x^n)$.

► Intégration et dérivation d'un développement limité

Théorème 42.1

Soit $f : I \rightarrow \mathbb{R}$ (avec $0 \in I$) une fonction dérivable en 0, telle que f' admet un développement limité d'ordre n au voisinage de 0, $f'(x) = a_0 + a_1x + \dots + a_nx^n + o(x^n)$, alors la fonction f admet au voisinage de 0 un développement limité d'ordre $n+1$ et :

$$f(x) = f(0) + a_0x + \frac{a_1}{2}x^2 + \dots + \frac{a_n}{n+1}x^{n+1} + o(x^{n+1}).$$

Théorème 42.2

Soit $f : I \rightarrow \mathbb{R}$ (avec $0 \in I$) une fonction dérivable en 0, telle que f admet un développement limité d'ordre n au voisinage de 0, $f(x) = a_0 + a_1x + \dots + a_nx^n + o(x^n)$, alors la fonction f' admet au voisinage de 0 un développement limité d'ordre $n-1$ et :

$$f'(x) = a_1 + 2a_2x + \dots + na_nx^{n-1} + o(x^{n-1}).$$

2. Utilisation des développements limités

Les développements limités servent à calculer des limites, à étudier des courbes. Pour les calculs, nous nous ramenons toujours à des développements limités en 0.

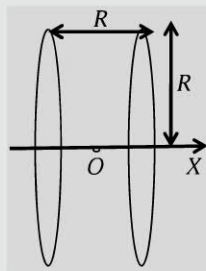
- Quand nous sommes au voisinage de x_0 , nous faisons alors le changement de variables $x = x_0 + h$. Ainsi nous écrivons des développements limités en 0 pour la variable h .
- Au voisinage de l'infini, nous utilisons le changement de variable $x = 1/t$ car quand x tend vers l'infini, alors $t = 1/x$ tend vers 0. Nous écrivons alors des développements limités en 0 pour la variable t . Ce changement de variable est utile pour étudier des fonctions qui tendent vers l'infini quand x tend vers l'infini (c'est-à-dire des fonctions admettant des branches infinies).

Application Les bobines de Helmholtz

Soient deux bobines plates identiques circulaires de N spires chacune et de même rayon R , placées dans deux plans parallèles, parcourues par le même courant I dans le même sens. La distance entre les centres des deux bobines est égale à R . Soit x l'abscisse d'un point M de l'axe (Ox) telle que $OM = x$. Le champ magnétique créé au point M par les bobines s'écrit :

$$\vec{B}(M) = B(x)\vec{e}_x = \left(f\left(\frac{R}{2} + x\right) + f\left(\frac{R}{2} - x\right)\right)\vec{e}_x \text{ avec } f(x) = \frac{\mu_0 N I R^2}{2} (R^2 + x^2)^{-\frac{3}{2}},$$

où μ_0 est la perméabilité du vide. Montrer que pour x voisin de 0, la fonction B peut s'écrire : $B(x) = 2f\left(\frac{R}{2}\right) + x^3 \epsilon(x)$ avec $\lim_{x \rightarrow 0} \epsilon(x) = 0$.



Posons $B_0 = \frac{\mu_0 N I}{2R}$, qui est une quantité homogène à un champ magnétique.

Nous avons : $f\left(\frac{R}{2} + x\right) = B_0 R^3 \left(R^2 + \left(\frac{R}{2} + x\right)^2\right)^{-\frac{3}{2}}$. En $x = 0$: $f\left(\frac{R}{2}\right) = B_0 R^3 \left(R^2 + \frac{R^2}{4}\right)^{-\frac{3}{2}} = B_0 \left(\frac{5}{4}\right)^{-\frac{3}{2}}$.



En développant à l'intérieur des () et en mettant en facteur, nous obtenons :

$$f\left(\frac{R}{2} + x\right) = B_0 \left(\frac{5}{4}\right)^{-\frac{3}{2}} \left(1 + \frac{4}{5} \times \frac{xR + x^2}{R^2}\right)^{-\frac{3}{2}} = f\left(\frac{R}{2}\right) \left(1 + \frac{4}{5} \times \frac{xR + x^2}{R^2}\right)^{-\frac{3}{2}}.$$



$(1+x)^\alpha = 1 + \alpha x + \frac{\alpha(\alpha-1)}{2!} x^2 + \dots + \frac{\alpha(\alpha-1)\dots(\alpha-n+1)}{n!} x^n + x^n \epsilon(x)$, avec $\lim_{x \rightarrow 0} \epsilon(x) = 0$.



Fiche 41

Or, la quantité $\frac{xR + x^2}{R^2}$ est très petite devant 1 car x est voisin de 0, et donc très petit devant la distance R . Donc, en appliquant le développement limité précédent à l'ordre 3, nous avons d'une part :

$$f\left(\frac{R}{2} + x\right) = f\left(\frac{R}{2}\right) \left(1 - \frac{3}{2} \left(\frac{4}{5} \times \frac{xR + x^2}{R^2}\right) + \frac{\frac{3}{2} \frac{5}{2}}{2} \left(\frac{4}{5} \times \frac{xR + x^2}{R^2}\right)^2 - \frac{\frac{3}{2} \frac{5}{2} \frac{7}{2}}{6} \left(\frac{4}{5} \times \frac{xR + x^2}{R^2}\right)^3\right) + f\left(\frac{R}{2}\right) \left(\frac{4}{5} \times \frac{xR + x^2}{R^2}\right)^3 \epsilon_1(x) \text{ avec } \lim_{x \rightarrow 0} \epsilon_1(x) = 0.$$

et d'autre part

$$f\left(\frac{R}{2} - x\right) = f\left(\frac{R}{2}\right) \left(1 - \frac{3}{2} \left(\frac{4}{5} \times \frac{-xR + x^2}{R^2}\right) + \frac{\frac{3}{2} \frac{5}{2}}{2} \left(\frac{4}{5} \times \frac{-xR + x^2}{R^2}\right)^2 - \frac{\frac{3}{2} \frac{5}{2} \frac{7}{2}}{6} \left(\frac{4}{5} \times \frac{-xR + x^2}{R^2}\right)^3\right) + f\left(\frac{R}{2}\right) \left(\frac{4}{5} \times \frac{-xR + x^2}{R^2}\right)^3 \epsilon_2(x) \text{ avec } \lim_{x \rightarrow 0} \epsilon_2(x) = 0.$$

$$\text{Donc, } B(x) = 2f\left(\frac{R}{2}\right) \left(1 - \frac{3}{2} \frac{4}{5} \frac{x^2}{R^2} + \frac{\frac{3}{2} \frac{5}{2}}{2} \left(\frac{4^2}{5^2} \times \frac{x^2}{R^2} + \frac{x^3}{R^3} \epsilon_3(x)\right)\right) \text{ avec } \lim_{x \rightarrow 0} \epsilon_3(x) = 0.$$

$$\text{Enfin, } B(x) = 2f\left(\frac{R}{2}\right) + \frac{x^3}{R^3} \epsilon_4(x) \text{ avec } \lim_{x \rightarrow 0} \epsilon_4(x) = 0.$$

Au voisinage du centre O du système constitué des deux bobines, le champ magnétique varie peu.

Pour s'entraîner (solutions p. 526)

42.1 Soit $n \in \mathbb{N}$, considérons la fonction f_n définie pour tout $x \in \mathbb{R}^*$ par :

$$f_n(x) = \frac{1}{x^{n+1}} \left(\exp(x) - \sum_{k=0}^n \frac{x^k}{k!} \right).$$

Donner le développement limité de la fonction exponentielle à l'ordre $2n + 1$ en 0. Exprimer $f_n(x)$, pour tout $x \in \mathbb{R}^*$, à l'aide de développement limité puis en déduire que f_n peut être prolongée par continuité en 0.

42.2 Nous rappelons que la fonction $\arcsin : x \rightarrow \arcsin(x)$, définie sur $[-1; 1]$ à valeurs dans $\left[-\frac{\pi}{2}; \frac{\pi}{2}\right]$, est continue sur $[-1; 1]$ et dérivable sur $] -1; 1[$.

De plus la fonction $\arcsin$ est la fonction réciproque de la fonction $\sin$, c'est-à-dire que pour tout $x \in [-1; 1]$ et tout $y \in \left[-\frac{\pi}{2}; \frac{\pi}{2}\right]$:

$$\sin \circ \arcsin(x) = x \quad \text{et} \quad \arcsin \circ \sin(y) = y.$$

Déterminer le développement limité à l'ordre 2 de $f : x \rightarrow (\arcsin(x))^2$ au voisinage de 0.

42.3 Nous considérons la fonction numérique f de la variable réelle x définie sur $[0; +\infty[$ par :

$$f(0) = 0 \quad \text{et} \quad f(x) = x^{1+\frac{1}{x}} \text{ si } x > 0.$$

Déterminer le développement limité à l'ordre 3 de f en 1. (Nous pourrions poser $x = 1 + t$).

42.4 Au XIX^e siècle, la loi de Rayleigh-Jeans donne la densité d'énergie du rayonnement du corps noir de longueur d'onde λ par $f(\lambda) = \frac{8\pi kT}{\lambda^4}$ avec T : température, k : constante de Boltzmann. Cette loi s'écarte des résultats expérimentaux pour les petites longueurs d'ondes. En 1900, la loi de Planck donne une nouvelle formule pour le rayonnement du corps noir : $f(\lambda) = \frac{8\pi hc \lambda^{-5}}{e^{\frac{hc}{\lambda kT}} - 1}$ avec h : constante de Planck, c : vitesse de la lumière. Montrer que pour les grandes longueurs d'ondes, nous retrouvons la loi de Rayleigh-Jeans.

Étude d'une fonction

Sauf dans la dernière section, nous considérons des fonctions à valeurs réelles. Pour le tracé des graphes des fonctions, nous travaillons dans le repère orthonormé du plan $(O, \vec{i}, \vec{j})$.

Définition 43.1

L'**ensemble de définition** D_f de la fonction f est l'ensemble des éléments x ($x \in \mathbb{R}$) ayant une image par f : $D_f = \{x \text{ tel que } f(x) \text{ existe}\}$.

1. Parité et périodicité

L'objectif est de réduire l'intervalle d'étude, i.e. l'intervalle sur lequel est étudiée la fonction f .

Définition 43.2

- La fonction f est **paire** si $\forall x \in D_f, (-x) \in D_f$ et $f(-x) = f(x)$.
Si f est paire, l'axe des ordonnées est axe de symétrie pour le graphe représentatif de f .
- La fonction f est **impaire** si $\forall x \in D_f, (-x) \in D_f$ et $f(-x) = -f(x)$.
Si f est impaire, l'origine est centre de symétrie pour le graphe représentatif de f .

Proposition 43.1

Si la fonction f est paire ou impaire, nous réduisons le domaine d'étude de la fonction à $D_f \cap \mathbb{R}_+$ et nous complétons la courbe représentative de la fonction en traçant le symétrique par rapport à l'axe des ordonnées ou l'origine, selon que f est paire ou impaire.

Définition 43.3

La fonction f est **périodique de période** $T \in \mathbb{R}_+^*$ ou encore **T -périodique** si $\forall x \in D_f, (x + T) \in D_f$ et $f(x + T) = f(x)$.
Si f est T -périodique, la courbe représentative est invariante pour toute translation de vecteur $kT \vec{i}$ ($k \in \mathbb{Z}$).

Proposition 43.2

Si la fonction f est T -périodique, nous pouvons étudier f sur un intervalle $[a; a + T[$, $a \in \mathbb{R}$, puis compléter la courbe obtenue par des translations de vecteurs $kT \vec{i}$ ($k \in \mathbb{Z}^*$).

2. Sens de variation et extremum local d'une fonction

Nous étudions le sens de variations et nous construisons ensuite le tableau de variation.

Proposition 43.3

Si f est une fonction dérivable sur un intervalle contenant x_0 et si la dérivée f' s'annule en x_0 en changeant de signe, la fonction f admet un **extremum local** en x_0 .

3. Branche infinie**Définition 43.4**

La courbe représentative $\mathcal{C}_f$ d'une fonction f admet une **branche infinie** lorsque OM tend vers l'infini où $M \in \mathcal{C}_f$. Lorsque $x \rightarrow +\infty$ ou $x \rightarrow -\infty$, ceci équivaut à $|f(x)| \rightarrow +\infty$.

Proposition 43.4

- Si f admet $+\infty$ ou $-\infty$ comme limite en x_0 ou comme limite à droite en x_0 ou comme limite à gauche en x_0 , la droite d'équation $x = x_0$ est **asymptote verticale** à la courbe représentative de f .
- Si f admet une limite l en $+\infty$ ou en $-\infty$, la droite d'équation $y = l$ est **asymptote horizontale** à la courbe représentative de f .

Proposition 43.5

Soit f admettant une limite infinie en $+\infty$ (resp. en $-\infty$).

- Si $\frac{f(x)}{x}$ admet une limite infinie en $+\infty$ (resp. en $-\infty$), la courbe $\mathcal{C}_f$ présente une **branche parabolique verticale**.
- Si $\frac{f(x)}{x}$ admet une limite finie a lorsque x tend vers $+\infty$ (resp. en $-\infty$) et si $f(x) - ax$ a une limite infinie, la courbe $\mathcal{C}_f$ présente une **branche parabolique de pente a** .
- Si $\frac{f(x)}{x}$ admet une limite finie a lorsque x tend vers $+\infty$ (resp. en $-\infty$) et s'il existe $b \in \mathbb{R}$ tel que $\lim_{x \rightarrow +\infty} [f(x) - (ax + b)] = 0$ (resp. $\lim_{x \rightarrow -\infty} [f(x) - (ax + b)] = 0$), la droite d'équation $y = ax + b$ est **asymptote oblique** à la courbe représentative de f .

4. Points remarquables**Définition 43.5**

Si, en un point de la courbe représentative d'une fonction continue, la concavité passe du type « convexe » au type « concave » (ou l'inverse), nous appelons ce point **point d'inflexion de la courbe**. Graphiquement, un point d'inflexion est un point où la tangente croise la courbe.



En un point d'inflexion de $\mathcal{C}_f$, la dérivée seconde de f , si elle existe, s'annule et change de signe. Ceci permet de vérifier si des points sont des points d'inflexion.

Définition 43.6

- Un point du graphe d'une fonction est un **point anguleux** si la dérivée à gauche de ce point n'est pas égale à la dérivée à droite de ce point et que l'une de ces dérivées au moins n'est pas infinie.

- Un point du graphe d'une fonction est un **point de rebroussement** si la dérivée à gauche de ce point n'est pas égale à la dérivée à droite de ce point et que ces deux dérivées sont infinies.



Une courbe représentative d'une fonction est un cas particulier de courbe paramétrée. Nous retrouverons toutes ces définitions lors de l'étude de celles-ci.



Fiche 63

5. Comparaisons entre fonctions

Définition 43.7

Soient $f : I \rightarrow \mathbb{K}$ et $g : I \rightarrow \mathbb{K}$, avec $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, et soit $x_0 \in I$.

- La fonction f est **négligeable** devant la fonction g au point x_0 , ce que nous notons $f(x) = o_{x_0}(g(x))$ si :

$$\forall \epsilon > 0, \exists \delta > 0, \text{ tel que } |x - x_0| < \delta \implies |f(x)| \leq \epsilon |g(x)|.$$

- La fonction f est **dominée** par la fonction g au point x_0 , ce que nous notons $f(x) = O_{x_0}(g(x))$ si :

$$\exists M > 0, \exists \delta > 0, \text{ tel que } |x - x_0| < \delta \implies |f(x)| \leq M |g(x)|.$$

- La fonction f est **équivalente** à la fonction g au point x_0 , ce que nous notons : $f(x) \sim_{x_0} g(x)$ si $f(x) - g(x) = o_{x_0}(g(x))$.

Application

La susceptibilité diélectrique

Les variations spatiales du champ électrique à l'échelle de l'atome sont négligeables. L'électron lié voit donc un champ électrique uniforme. Nous souhaitons décrire alors sa réponse à un champ électrique variant sinusoidalement dans le temps. C'est un problème d'oscillateur harmonique amorti en régime sinusoidal forcé. De l'étude, il résulte un moment dipolaire volumique de représentation complexe :

$$\vec{P} = \epsilon_0 \underline{\chi} \vec{E}$$

où

- $\vec{E}$ est le vecteur champ électrique,
- $\underline{\chi}$ la susceptibilité diélectrique avec : $\underline{\chi} = \chi_1 - i\chi_2$ avec

$$\chi_1 = \frac{1 - \left(\frac{w}{w_0}\right)^2}{\left(1 - \left(\frac{w}{w_0}\right)^2\right)^2 + \left(\frac{w}{10w_0}\right)^2} \quad \text{et} \quad \chi_2 = \frac{\frac{w}{10w_0}}{\left(1 - \left(\frac{w}{w_0}\right)^2\right)^2 + \left(\frac{w}{10w_0}\right)^2}$$

où w_0 est la pulsation propre, i est l'imaginaire pur tel que $i^2 = -1$ et ϵ_0 est la permittivité du vide.

Nous posons : $x = \frac{w}{w_0}$ avec $x \geq 0$. Les deux dernières expressions s'écrivent alors :

$$\chi_1(x) = \frac{1 - x^2}{(1 - x^2)^2 + \frac{x^2}{100}} \quad \text{et} \quad \chi_2(x) = \frac{\frac{x}{10}}{(1 - x^2)^2 + \frac{x^2}{100}}.$$

Étudier les variations de ces deux fonctions sur $[0; +\infty[$.

Les dénominateurs ne s'annulent pas pour $x \in [0; +\infty[$, le domaine de définition de χ_1 et χ_2 est donc $[0; +\infty[$. Les fonctions χ_1 et χ_2 sont dérivables sur leur domaine de définition, et :

$$(\chi_1)'(x) = \frac{-2x\left((1-x^2)^2 + \frac{x^2}{100}\right) - (1-x^2)\left(-4x(1-x^2) + \frac{2x}{100}\right)}{\left((1-x^2)^2 + \frac{x^2}{100}\right)^2} = \frac{x(100x^4 - 200x^2 + 99)}{50\left((1-x^2)^2 + \frac{x^2}{100}\right)^2}.$$

Étudions maintenant le signe de $P(x) = 100x^4 - 200x^2 + 99$. Pour cela, nous devons résoudre l'équation $100x^4 - 200x^2 + 99 = 0$. Nous posons $X = x^2$ et ainsi nous sommes ramenés à étudier : $100X^2 - 200X + 99 = 0$. Nous obtenons deux racines $X_1 = \frac{11}{10}$ et $X_2 = \frac{9}{10}$ d'où les quatre racines de P sont $x_1 = \sqrt{\frac{11}{10}}$, $x_2 = \sqrt{\frac{9}{10}}$, $x_3 = -x_1$, $x_4 = -x_2$. Donc, χ_1' s'écrit alors :

$$(\chi_1)'(x) = \frac{2x(x - \sqrt{\frac{11}{10}})(x - \sqrt{\frac{9}{10}})(x + \sqrt{\frac{11}{10}})(x + \sqrt{\frac{9}{10}})}{\left((1-x^2)^2 + \frac{x^2}{100}\right)^2}.$$

Donc $(\chi_1)'(x)$ est strictement positive pour $x \in \left]0; \sqrt{\frac{9}{10}}\right[\cup \left]\sqrt{\frac{11}{10}}; +\infty\right[$ et strictement négative pour $x \in \left]\sqrt{\frac{9}{10}}; \sqrt{\frac{11}{10}}\right[$. Donc χ_1 est strictement croissante sur $\left]0; \sqrt{\frac{9}{10}}\right[\cup \left]\sqrt{\frac{11}{10}}; +\infty\right[$ et strictement décroissante sur $\left]\sqrt{\frac{9}{10}}; \sqrt{\frac{11}{10}}\right[$. Maintenant, étudions χ_2 .

$$(\chi_2(x))' = \frac{\frac{1}{10}\left((1-x^2)^2 + \frac{x^2}{100}\right) - \frac{x}{10}\left(-4x(1-x^2) + \frac{x}{50}\right)}{\left((1-x^2)^2 + \frac{x^2}{100}\right)^2} = \frac{-300x^4 + 199x^2 + 100}{1000\left((1-x^2)^2 + \frac{x^2}{100}\right)^2}.$$

Étudions maintenant le signe de $Q(x) = -300x^4 + 199x^2 + 100$. Nous posons $X = x^2$ et ainsi nous sommes ramenés à étudier : $-300X^2 + 199X + 100 = 0$. Nous obtenons, après calculs : $X_1 = \frac{199 + \sqrt{159601}}{600}$ et $X_2 = \frac{199 - \sqrt{159601}}{600} < 0$. Donc $x_1 = \sqrt{\frac{199 + \sqrt{159601}}{600}}$ et $x_2 = -x_1$ sont les seules racines réelles de Q donc $(\chi_2(x))'$ est positive pour $x \in]0; x_1[$, et négative pour $x \in]x_1; +\infty[$. Donc χ_2 est strictement croissante sur $x \in]0; x_1[$ et strictement décroissante sur $x \in]x_1; +\infty[$.

Pour s'entraîner (solutions p. 526)

43.1 Étudier le domaine de définition, les symétries s'il y en a, le sens de variations et les éventuelles branches infinies des fonctions suivantes :

1. $f(x) = \frac{x^2 + x}{x^2 - 4}$;

2. $f(x) = \cos(3x) \sin(x)$;

3. $f(x) = xe^{-x \ln(2)}$.

43.2 1. Si une fonction est impaire et bijective, que dire de sa fonction réciproque f^{-1} ?

2. Les fonctions arctan et argh sont-elles impaires ?

Intégrales définies sur segment

1. Intégrale d'une fonction en escalier

Dans cette Fiche, $[a; b]$ désigne un intervalle de $\mathbb{R}$, avec $a < b$.

► Fonction en escalier

Définition 44.1

- Nous appelons **subdivision** σ de $[a; b]$ toute suite finie de points $(x_0 = a < x_1 < \dots < x_n = b)$.
- L'ensemble $\{x_0, \dots, x_n\}$ est appelé le **support de la subdivision**.
- La quantité $\max_{0 \leq i \leq n-1} (x_{i+1} - x_i)$ est appelée le **pas d'une subdivision** $(x_i)_{0 \leq i \leq n}$.

Définition 44.2

Soit f une fonction de $[a; b]$ dans $\mathbb{R}$. Nous disons que la **fonction f est en escalier sur $[a; b]$** s'il existe une subdivision $(x_i)_{0 \leq i \leq n}$ de $[a; b]$ et n réels $\lambda_0, \dots, \lambda_{n-1}$ tels que :

$$\forall i = 0, \dots, n-1, \quad \forall t \in]x_i; x_{i+1}[, \quad f(t) = \lambda_i.$$

Nous disons que la subdivision $(x_i)_{0 \leq i \leq n}$ est adaptée à f . Nous notons $\mathcal{E}([a; b], \mathbb{R})$ l'ensemble des fonctions en escalier sur $[a; b]$.

Définition 44.3

Soit I un intervalle de $\mathbb{R}$ d'intérieur non vide. Soit ϕ une fonction de I dans $\mathbb{R}$. Nous disons que la **fonction ϕ est en escalier sur I** s'il existe un intervalle $[a; b]$ de I tel que :

1. La fonction ϕ est nulle en dehors de $[a; b]$.
2. La restriction de la fonction ϕ à $[a; b]$, notée ψ , est en escalier sur $[a; b]$. Une subdivision de l'intervalle $[a; b]$ adaptée à ψ est encore dite adaptée à ϕ .

► Intégrale des fonctions en escalier

Définition 44.4

Soient $\phi : [a; b] \rightarrow \mathbb{R}$ une fonction en escalier et $\sigma = (x_i)_{0 \leq i \leq n}$ une subdivision adaptée. Nous supposons que : $\forall i \in \{0; \dots; n-1\}, \quad \exists \lambda_i \in \mathbb{R}, \quad \forall t \in]x_i; x_{i+1}[, \quad \phi(t) = \lambda_i$. Le nombre réel $I(\phi) = \sum_{i=0}^{n-1} (x_{i+1} - x_i) \lambda_i$ est appelé **intégrale** de ϕ et est noté $\int_a^b \phi(t) dt$.

- L'intégrale de ϕ ne dépend pas de la subdivision adaptée à ϕ .
- Si ϕ est une fonction constante et est égale à λ sur $[a; b]$, alors $\int_a^b \phi(t) dt = (b-a)\lambda$.



- Soit ϕ un élément de $\mathcal{E}([a; b], \mathbb{R})$. Soit c un élément de $]a; b[$. Alors les restrictions de ϕ à $[a; c]$ et $[c; b]$ sont en escalier et nous avons la relation suivante : $\int_a^b \phi(t)dt = \int_a^c \phi(t)dt + \int_c^b \phi(t)dt$.

2. Intégrale d'une fonction continue par morceaux

► Fonction continue par morceaux

Définition 44.5

Nous disons qu'une application $f : [a; b] \rightarrow \mathbb{R}$ est **continue par morceaux** sur $[a; b]$ s'il existe une subdivision $\sigma = (x_i)_{0 \leq i \leq n}$ de $[a; b]$ (dite adaptée à f , ou encore subordonnée à f) telle que, pour tout i de $\{0; \dots; n-1\}$ la restriction f_i de f à $]x_i; x_{i+1}[$ est continue et celle-ci est prolongeable par continuité aux points x_i et x_{i+1} . Nous notons $\mathcal{C}_m([a; b], \mathbb{R})$ l'ensemble des fonctions continues par morceaux sur $[a; b]$.

► Intégrale d'une fonction continue par morceaux

Proposition 44.1 Approximation par une fonction en escalier

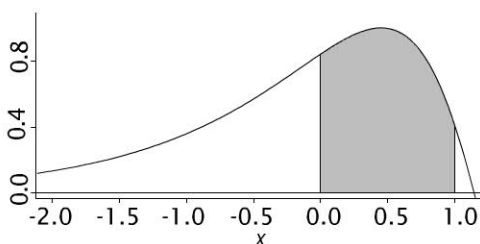
Soit f une fonction continue par morceaux sur $[a; b]$. Pour tout réel $\varepsilon > 0$, il existe deux fonctions en escalier ϕ et ψ telles que : pour tout x de $[a; b]$, $\phi(x) \leq f(x) \leq \psi(x)$ et $0 \leq \psi(x) - \phi(x) \leq \varepsilon$, relations qui se notent $\phi \leq f \leq \psi$ et $0 \leq \psi - \phi \leq \varepsilon$.

Définition 44.6

Soit f une fonction continue par morceaux sur $[a; b]$. Il existe un unique réel $I(f)$ tel que, pour toutes fonctions ϕ et ψ en escalier sur $[a; b]$ vérifiant l'inégalité suivante $\phi \leq f \leq \psi$, nous ayons : $I(\phi) \leq I(f) \leq I(\psi)$. Ce nombre s'appelle l'**intégrale de la fonction f sur $[a; b]$** . Il se note $\int_a^b f(t)dt$. La fonction f est alors dite **intégrable** sur le segment $[a; b]$.



- L'intégrale dépend de la fonction f et des deux réels a et b mais pas de la variable d'intégration, qui nous le rappelons, est une variable muette.
- Pour $a < b$, nous avons : $\int_b^a f(t)dt = - \int_a^b f(t)dt$.
- L'intégrale $\int_a^b f(t)dt$ correspond à la surface délimitée par la représentation graphique de la fonction f et par les trois droites d'équation $x = a$, $x = b$ et $y = 0$.



L'intégrale $\int_0^1 f(t)dt$ est égale à l'aire sous la courbe entre les abscisses 0 et 1.

► Propriétés de l'intégrale

Soient f et g deux fonctions continues par morceaux sur $[a; b]$, avec $a < b$.

• Linéarité de l'intégrale

Pour tous α et β des réels, nous avons : $\int_a^b (\alpha f(t) + \beta g(t)) dt = \alpha \int_a^b f(t) dt + \beta \int_a^b g(t) dt$.

• Positivité et croissance de l'intégrale

- Si la fonction f est positive ou nulle sur l'intervalle $[a; b]$, alors $\int_a^b f(t) dt \geq 0$.
- Si l'inégalité suivante est vérifiée $f \leq g$ sur l'intervalle $[a; b]$, alors nous avons $\int_a^b f(t) dt \leq \int_a^b g(t) dt$.

• Inégalité de la moyenne

Nous avons l'inégalité suivante : $\left| \int_a^b f(t) \times g(t) dt \right| \leq \sup_{t \in [a; b]} |f(t)| \times \int_a^b |g(t)| dt$. En particulier nous avons : $\left| \int_a^b f(t) dt \right| \leq \int_a^b |f(t)| dt \leq \sup_{t \in [a; b]} |f(t)| \times (b - a)$.

Définition 44.7

Soit f une fonction continue par morceaux sur $[a; b]$ avec $a < b$. La quantité $\frac{1}{b-a} \int_a^b f(t) dt$ est appelée la **valeur moyenne** de la fonction f sur $[a; b]$.

• Inégalité de Cauchy-Schwarz

Nous avons l'inégalité suivante : $\left(\int_a^b f(t) \times g(t) dt \right)^2 \leq \int_a^b f^2(t) dt \times \int_a^b g^2(t) dt$.



Si les fonctions f et g sont continues, alors il y a égalité si et seulement si les deux fonctions f et g sont proportionnelles, i.e. $\exists \lambda \in \mathbb{R}, f = \lambda g$.

• Relation de Chasles

Pour tout réel $c \in]a; b[$, nous avons : $\int_a^b f(t) dt = \int_a^c f(t) dt + \int_c^b f(t) dt$.

3. Un exemple de calcul approché des intégrales : les sommes de Riemann

Définition 44.8

Soit f une fonction de $[a; b]$ dans $\mathbb{R}$. Soit $\sigma = (x_i)_{0 \leq i \leq n}$ une subdivision de l'intervalle $[a; b]$. Pour chaque i de $\{0; \dots; n-1\}$, nous choisissons un réel ζ_i dans $[x_i; x_{i+1}]$. La quantité $R_S(f) = \sum_{i=0}^{n-1} (x_{i+1} - x_i) f(\zeta_i)$ est appelée **somme de Riemann** de f associée à la subdivision σ (et au choix des abscisses ζ_i).



Un cas particulier classique est celui où la subdivision σ est régulière, c'est-à-dire formée en divisant l'intervalle $[a; b]$ en n sous-intervalles de même longueur $h = \frac{b-a}{n}$. Pour tout i de $\{0; \dots; n\}$, nous avons alors $x_i = a + i \times h$. Les sommes de Riemann ainsi construites

s'écrivent $R_S(f) = \frac{b-a}{n} \sum_{i=0}^{n-1} f(\zeta_i)$. Nous choisissons souvent $\zeta_i = x_i$, ou $\zeta_i = x_{i+1}$, ou encore

$$\zeta_i = \frac{1}{2} (x_i + x_{i+1}).$$

Proposition 44.2 Convergence des sommes de Riemann

Soit f une fonction continue par morceaux sur l'intervalle $[a; b]$. Soit ε un réel strictement positif. Alors il existe un réel $\delta > 0$ tel que pour toute subdivision σ de l'intervalle $[a; b]$ de pas inférieur ou égal à δ nous ayons : $\left| \int_a^b f(t)dt - R_S(f) \right| \leq \varepsilon$. Ceci signifie que la somme de Riemann converge vers $\int_a^b f(t)dt$ lorsque le pas de la subdivision tend vers 0.

Proposition 44.3 Cas particulier

Soit f une fonction continue par morceaux sur l'intervalle $[a; b]$. Alors nous avons l'égalité suivante :

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{i=0}^{n-1} f\left(a + i \frac{b-a}{n}\right) = \frac{1}{b-a} \int_a^b f(t)dt.$$



- Dans ce résultat, nous pouvons remplacer $\sum_{i=0}^{n-1}$ par $\sum_{i=0}^n$ ou par $\sum_{i=1}^n$, cela ne change rien.
- La proposition précédente permet de calculer des limites de suites dont le terme général u_n peut être interprété comme une somme de Riemann. Il est recommandé de comparer u_n avec la forme générale d'une somme de Riemann, pour éviter toute erreur sur l'intervalle $[a; b]$ et sur la fonction f .

Application**1. Utilisation d'une somme de Riemann**

Soit $u_n = \sum_{k=n+1}^{2n} \frac{1}{k}$, avec $n \in \mathbb{N}^*$. Calculer $\lim_{n \rightarrow +\infty} u_n$.

2. Utilisation de l'inégalité de Cauchy-Schwarz

Soit f une fonction continue sur $[0; 1]$ à valeurs positives ou nulles. Posons $I_n = \int_0^1 f^n(t)dt$. Montrer que : $\forall n \in \mathbb{N}^*, I_n^2 \leq I_{n-1}I_{n+1}$.

1. Nous posons $k = p + n$. Ainsi, $u_n = \sum_{p=1}^n \frac{1}{p+n} = \frac{1}{n} \sum_{p=1}^n \frac{1}{1 + \frac{p}{n}} = \frac{1}{n} \sum_{p=1}^n f\left(\frac{p}{n}\right)$ où $f(t) = \frac{1}{1+t}$ qui est une fonction continue sur $[0; 1]$. Ainsi u_n est une somme de Riemann. Donc $\lim_{n \rightarrow +\infty} u_n = \int_0^1 \frac{1}{1+t} dt = [\ln(1+t)]_0^1 = \ln(2)$. La suite (u_n) converge vers $\ln(2)$.

2. Pour tout $n \in \mathbb{N}^*$ et $t \in [0; 1]$, nous avons $f^n(t) = f^{\frac{n-1}{2}}(t)f^{\frac{n+1}{2}}(t)$. L'inégalité de Cauchy-Schwarz donne : $\left(\int_0^1 f^n(t)dt\right)^2 \leq \int_0^1 \left(f^{\frac{n-1}{2}}(t)\right)^2 dt \times \int_0^1 \left(f^{\frac{n+1}{2}}(t)\right)^2 dt$, c'est-à-dire $I_n^2 \leq I_{n-1}I_{n+1}$.

Pour s'entraîner
(solutions p. 527)

44.1 Soit $v_n = \sum_{k=1}^n \frac{n}{k^2 + n^2}$, avec $n \in \mathbb{N}^*$. Calculer $\lim_{n \rightarrow +\infty} v_n$.

44.2 Montrer que la fonction f définie par : $f(x) = \begin{cases} x & \text{si } x < 1 \\ -x & \text{si } x \geq 1 \end{cases}$ est continue par morceaux sur $[0; 2]$.

Primitives et intégrales d'une fonction continue

Dans cette Fiche, I désigne un intervalle d'intérieur non vide de $\mathbb{R}$.

1. Le théorème fondamental

Définition 45.1 Primitive sur un intervalle

Soit $f : I \rightarrow \mathbb{R}$ une fonction. Une fonction $F : I \rightarrow \mathbb{R}$ est une **primitive de la fonction** f , si F est dérivable sur I et si pour tout $x \in I$, nous avons : $F'(x) = f(x)$.

Proposition 45.1

Relation entre les primitives d'une même fonction Soit $f : I \rightarrow \mathbb{R}$ admettant une primitive F sur I . Une fonction $G : I \rightarrow \mathbb{R}$ est une primitive de f sur I si et seulement s'il existe une constante λ telle que, pour tout x de I , $G(x) = F(x) + \lambda$.

Proposition 45.2

Primitive prenant une valeur donnée en un point donné

Soit $f : I \rightarrow \mathbb{R}$ une fonction admettant une primitive F sur I . Soient $a \in I$ et λ un réel. Il y a une seule primitive G de f telle que $G(a) = \lambda$. Elle est donnée par $G = F - F(a) + \lambda$. En particulier, $H = F - F(a)$ est l'unique primitive de la fonction f sur I qui s'annule au point a .

Théorème 45.1 Le théorème « fondamental »

Si $f : I \rightarrow \mathbb{R}$ est une fonction continue sur I , alors la fonction f admet des primitives sur I .

Proposition 45.3 Expression d'une intégrale à l'aide d'une primitive quelconque

Soit $f : I \rightarrow \mathbb{R}$ une fonction continue sur I . Pour toute primitive F de f , nous avons : $\forall (a; b) \in I^2, \int_a^b f(t)dt = [F(t)]_a^b = F(b) - F(a)$.

2. Formule de Taylor avec reste intégral

Proposition 45.4

Soient une fonction f de classe $\mathcal{C}^{n+1}$ sur I et x_0 et x des points appartenant à I . Nous avons alors : $f(x) = P_n(x) + R_n(x)$ où $P_n(x) = f(x_0) + \frac{(x-x_0)}{1!}f'(x_0) + \dots + \frac{(x-x_0)^n}{n!}f^{(n)}(x_0)$ et $R_n(x) = \int_{x_0}^x \frac{(x-t)^n}{n!}f^{(n+1)}(t)dt$ est le **reste intégral d'ordre n** . Cette écriture est appelée l'**approximation de Taylor avec reste intégral** de la fonction f à l'ordre n .

3. Méthodes de calcul des intégrales

Si l'intégrale cherchée ne peut pas être obtenue immédiatement par utilisation d'une primitive usuelle, nous chercherons à transformer l'intégrale initiale en une ou plusieurs autres que nous savons calculer.

► Intégration par parties

Théorème 45.2

Soient $f : I \rightarrow \mathbb{R}$ et $g : I \rightarrow \mathbb{R}$ deux fonctions de classe $\mathcal{C}^1$ sur I et a et b des réels de I . Nous avons : $\int_a^b f(t)g'(t)dt = [f(t)g(t)]_a^b - \int_a^b f'(t)g(t)dt$.

► Changement de variable

Proposition 45.5

Soient I et J deux intervalles de $\mathbb{R}$ d'intérieur non vide. Soit $f : I \rightarrow \mathbb{R}$ une fonction continue sur I . Soit $\phi : J \rightarrow \mathbb{R}$ une fonction de classe $\mathcal{C}^1$ sur J , telle que $\phi(J) \subset I$. Alors, pour tous points a et b de l'intervalle J , nous avons l'égalité suivante : $\int_{\phi(a)}^{\phi(b)} f(x)dx = \int_a^b (f \circ \phi)(t)\phi'(t)dt$.

4. Compléments sur le calcul des primitives

► Primitives de $\sin^p(t) \times \cos^q(t)$

Si nous souhaitons calculer $\int \sin^p(t) \cos^q(t)dt$, avec p et q deux entiers naturels, tout dépend de la parité de p et de q .

- Si p est impair, nous posons alors $x = \cos(t)$ et donc $dx = -\sin(t)dt$.
- Si q est impair, nous posons alors $x = \sin(t)$ et donc $dx = \cos(t)dt$.
- Si p et q sont pairs, nous linéarisons.

► Fractions rationnelles

Étudions la primitive suivante : $\int \frac{ax+b}{(x^2+px+q)^n} dx$ où $n \in \mathbb{N}^*$. Posons : $\Delta = p^2 - 4q$.



Point méthode

Si $\Delta \geq 0$ alors nous décomposons en éléments simples.

Si $\Delta < 0$, nous faisons apparaître au numérateur la dérivée du dénominateur.



Fiche 13

- Cas $n = 1$. Nous avons $\frac{ax+b}{x^2+px+q} = \frac{\frac{a}{2}(2x+p) - \frac{ap}{2} + b}{x^2+px+q}$. Donc $\int \frac{ax+b}{x^2+px+q} dx = \frac{a}{2} \ln |x^2+px+q| + \left(\frac{-ap}{2} + b\right) \int \frac{dx}{x^2+px+q}$. Maintenant traitons : $\int \frac{dx}{x^2+px+q}$. Pour cela, faisons le changement de variables $u = x + \frac{p}{2}$ et posons $\gamma^2 = q - \frac{p^2}{4}$: $\int \frac{dx}{x^2+px+q} = \int \frac{du}{u^2+\gamma^2}$. Donc $\int \frac{dx}{x^2+px+q} = \frac{1}{\gamma} \arctan\left(\frac{x+\frac{p}{2}}{\gamma}\right) + K$, où K est une constante.

- Cas $n > 1$. Nous avons $\int \frac{ax+b}{(x^2+px+q)^n} dx = \frac{a}{2} \int \frac{2x+p}{(x^2+px+q)^n} dx + \left(\frac{-ap}{2} + b\right) \int \frac{dx}{(x^2+px+q)^n} = -\frac{a}{2(n-1)(x^2+px+q)^{n-1}} + \left(\frac{-ap}{2} + b\right) \int \frac{dx}{(x^2+px+q)^n}$. Traitons maintenant : $\int \frac{dx}{(x^2+px+q)^n}$. Pour cela, faisons le changement de variable $v = \frac{x}{\gamma} + \frac{p}{2\gamma}$, $\gamma^2 = q - \frac{p^2}{4}$: $\int \frac{dx}{(x^2+px+q)^n} = \frac{1}{\gamma^{2n-1}} \int \frac{dv}{(v^2+1)^n}$. Nous calculons $I_n = \int \frac{dv}{(v^2+1)^n}$ avec une relation de récurrence. Nous obtenons pour tout $n \geq 1$: $I_{n+1} = \frac{2n-1}{2n} I_n + \frac{v}{2n(1+v^2)^n}$.

► Règles de Bioche

Nous considérons ici des expressions rationnelles $R(\sin t, \cos t, \tan t)$, c'est-à-dire formées par des sommes, des produits et des puissances entières de fonctions trigonométriques. Les règles de Bioche consistent à proposer un changement de variable quand l'expression $R(\sin t, \cos t, \tan t)dt$ est invariante dans une certaine transformation. Ces changements de variable conduisent à une fraction rationnelle.

- Nous avons « l'invariant du cosinus » si $R(\sin t, \cos t, \tan t)dt$ est inchangé dans $t \rightarrow -t$. Dans ce cas, nous pouvons faire le changement de variable $x = \cos t$.
- Nous avons « l'invariant du sinus » si $R(\sin t, \cos t, \tan t)dt$ est inchangé dans $t \rightarrow \pi - t$. Dans ce cas, nous pouvons faire le changement de variable $x = \sin t$.
- Nous avons « l'invariant de la tangente » si $R(\sin t, \cos t, \tan t)dt$ est inchangé dans $t \rightarrow t + \pi$. Dans ce cas, nous pouvons faire le changement de variable $x = \tan t$.
- Et si rien ne marche, nous pouvons toujours effectuer le changement de variable $x = \tan\left(\frac{t}{2}\right)$, qui nous ramène à une fraction rationnelle. L'inconvénient est que le degré est doublé.

Exemples

Calculer $I_1 = \int \frac{dx}{2+\cos(x)}$ et $I_2 = \int \frac{dx}{2+\sin(x)}$.

Posons $t = \tan\left(\frac{x}{2}\right)$, nous avons alors $I_1 = \int \frac{2dt}{(1+t^2)\left(\frac{1-t^2}{1+t^2}+2\right)} = 2 \int \frac{dt}{3+t^2} = \frac{2}{3} \int \frac{dt}{1+\left(\frac{t}{\sqrt{3}}\right)^2} = \frac{2\sqrt{3}}{3} \arctan\left(\frac{t}{\sqrt{3}}\right) + \text{constante} = \frac{2\sqrt{3}}{3} \arctan\left(\frac{1}{\sqrt{3}} \tan\left(\frac{x}{2}\right)\right) + \text{constante}$. De même, $I_2 = \int \frac{dt}{t^2+t+1} = \int \frac{ds}{s^2+\frac{3}{4}} = \frac{2}{\sqrt{3}} \arctan\left(\frac{2s}{\sqrt{3}}\right) + \text{constante}$ avec $s = t + \frac{1}{2}$. Donc, $I_2 = \frac{2}{\sqrt{3}} \arctan\left(\frac{2t+1}{\sqrt{3}}\right) + \text{constante} = \frac{2}{\sqrt{3}} \arctan\left(\frac{2\tan\left(\frac{x}{2}\right)+1}{\sqrt{3}}\right) + \text{constante}$.

► Fraction rationnelle en e^x , $\text{sh}(x)$, $\text{ch}(x)$

Nous pouvons faire comme changement de variables : $u = \text{th}\left(\frac{x}{2}\right)$ ou $u = e^x$.

Exemple

Calculer $I_1 = \int_0^{\ln(4)} \frac{3dx}{5\text{sh}(x)-4\text{ch}(x)}$.

$I_1 = \int_0^{\ln(4)} \frac{3dx}{5\left(\frac{e^x-e^{-x}}{2}\right)-4\left(\frac{e^x+e^{-x}}{2}\right)} = \int_0^{\ln(4)} \frac{6e^x dx}{e^{2x}-9}$. Posons $u = e^x$. Alors nous avons

$$I_1 = \int_1^4 \frac{6du}{u^2-9} = \int_1^4 \left(\frac{1}{u-3} - \frac{1}{u+3} \right) du = [\ln|u-3| - \ln|u+3|]_1^4 = \ln\left(\frac{2}{7}\right).$$

Application

L'énergie proton-électron

L'étude spectroscopique des émissions lumineuses produites par un gaz d'atomes d'hydrogène induit à penser que l'énergie mécanique totale du système lié constitué par le proton et l'électron ne peut prendre que des valeurs discrètes. Nous introduisons alors dans le cas du modèle de Bohr la quantité :

$$I = \frac{L}{\pi} \int_0^\pi \frac{\xi^2 \sin^2(\theta)}{(1 + \xi \cos(\theta))^2} d\theta,$$

où L représente le moment cinétique et $0 < \xi < 1$ l'excentricité. Calculer I .



Posons $u(\theta) = \xi \sin(\theta)$ et $v(\theta) = \frac{1}{1 + \xi \cos(\theta)}$, $u'(\theta) = \xi \cos(\theta)$ et $v'(\theta) = \frac{\xi \sin(\theta)}{(1 + \xi \cos(\theta))^2}$. $u, v \in \mathcal{C}^1([0; \pi])$.

Nous avons alors : $\frac{\pi}{L}I = \left[\frac{\xi \sin(\theta)}{1 + \xi \cos(\theta)} \right]_0^\pi - \int_0^\pi \frac{\xi \cos(\theta)}{1 + \xi \cos(\theta)} d\theta = -\xi \int_0^\pi \frac{\cos(\theta)}{1 + \xi \cos(\theta)} d\theta.$

Or, $\int_0^\pi d\theta = \pi = \int_0^\pi \frac{(1 + \xi \cos(\theta))}{1 + \xi \cos(\theta)} d\theta$, donc $\pi - \int_0^\pi \frac{d\theta}{1 + \xi \cos(\theta)} = \int_0^\pi \frac{\xi \cos(\theta)}{1 + \xi \cos(\theta)} d\theta$. Calculons : $\int_0^\pi \frac{d\theta}{1 + \xi \cos(\theta)}.$



Pour pouvoir procéder à un changement de variable en $t = \tan\left(\frac{\theta}{2}\right)$, nous devons nous placer sur l'intervalle $[0; \pi - \varepsilon]$ où $\varepsilon \in]0; \pi[$, car $\tan$ n'est pas définie en $\frac{\pi}{2}$.

Posons $t = \tan\left(\frac{\theta}{2}\right)$, donc $dt = \frac{(1 + t^2)d\theta}{2}$ et $\cos(\theta) = \frac{1 - t^2}{1 + t^2}$. En posant $I_\varepsilon = \int_0^{\pi-\varepsilon} \frac{d\theta}{1 + \xi \cos(\theta)}$, nous avons

$$I_\varepsilon = \int_0^{\pi-\varepsilon} \frac{d\theta}{1 + \xi \cos(\theta)} = \int_0^{\tan(\frac{\pi-\varepsilon}{2})} \frac{2dt}{1 + t^2 + \xi(1 - t^2)} = \frac{2}{1 - \xi} \int_0^{\tan(\frac{\pi-\varepsilon}{2})} \frac{dt}{t^2 + \frac{1+\xi}{1-\xi}}.$$

Or, $\int \frac{dx}{x^2 + a^2} = \frac{1}{a} \arctan\left(\frac{x}{a}\right)$. Donc $\int_0^{\pi-\varepsilon} \frac{d\theta}{1 + \xi \cos(\theta)} = \left[\frac{2}{1-\xi} \sqrt{\frac{1-\xi}{1+\xi}} \arctan\left(\sqrt{\frac{1-\xi}{1+\xi}} t\right) \right]_0^{\tan(\frac{\pi-\varepsilon}{2})} = \frac{2}{\sqrt{1-\xi^2}} \times \arctan\left(\sqrt{\frac{1-\xi}{1+\xi}} \tan\left(\frac{\pi-\varepsilon}{2}\right)\right)$. Cette expression admet une limite lorsque ε tend vers 0 puisqu'alors $\tan\left(\frac{\pi-\varepsilon}{2}\right)$ tend vers $+\infty$ et de ce fait, comme $\sqrt{\frac{1-\xi}{1+\xi}} > 0$, le terme $\arctan\left(\sqrt{\frac{1-\xi}{1+\xi}} \tan\left(\frac{\pi-\varepsilon}{2}\right)\right)$ tend vers $\frac{\pi}{2}$. Donc, nous en déduisons, $\frac{\pi}{L}I = -\int_0^\pi \frac{\xi \cos(\theta)d\theta}{1 + \xi \cos(\theta)} = \lim_{\varepsilon \rightarrow 0} I_\varepsilon - \pi = \frac{2}{\sqrt{1-\xi^2}} \frac{\pi}{2} - \pi$.

En conclusion, $I = L\left(\frac{1}{\sqrt{1-\xi^2}} - 1\right)$.

Pour s'entraîner (solutions p. 527)

45.1 Déterminer une primitive de $f(x) = \frac{1}{\sin(x)}$, pour $x \in]0; \pi[$.

45.2 Déterminer une primitive de $f(x) = \frac{\sqrt{x^{10} + 1}}{x}$, pour $x \in]0; +\infty[$.

45.3 Déterminer une primitive de

1. $f(x) = \frac{1}{x^4 + 4}$;
2. $f(x) = \frac{5x - 4}{2x^2 + x - 1}$;
3. $f(x) = \frac{2x^2 - x + 4}{x^3 + 4x}$.

45.4 Déterminer une primitive de $f(x) = (x + 1) \arctan\left(\sqrt{\frac{x+3}{x+1}}\right)$, pour $x \in]-1; +\infty[$.

45.5 Déterminer une primitive de $f(x) = \frac{1 + x + x^2}{\sqrt{1 - x^2}}$.

45.6 Déterminer une primitive de $f(x) = \frac{1}{5 \operatorname{ch}(x) + 3 \operatorname{sh}(x) + 4}$.

45.7 Déterminer une primitive de $f(x) = \frac{1 - \cos\left(\frac{x}{3}\right)}{\sin\left(\frac{x}{2}\right)}$.

1. Définitions

Définition 46.1 Sommes partielles d'une série

Soit $(u_n)_{n \geq 0}$ une suite d'éléments de $\mathbb{K}$ ($= \mathbb{R}$ ou $\mathbb{C}$) et $N \in \mathbb{N}$. Nous appelons **somme partielle** d'indice N de la série $\sum_{n \geq 0} u_n$ la quantité $S_N = \sum_{n=0}^N u_n$.

Définition 46.2 Convergence ou divergence d'une série

Soit $(u_n)_{n \geq 0}$ une suite d'éléments de $\mathbb{K}$ ($= \mathbb{R}$ ou $\mathbb{C}$). Nous disons que la **série** $\sum_{n \geq 0} u_n$ ou la série de terme général u_n est **convergente** si la suite S_N de ses sommes partielles est convergente. Dans le cas contraire, nous disons que la **série** $\sum_{n \geq 0} u_n$ est **divergente**.

Définition 46.3 Somme d'une série convergente

Soit $\sum_{n \geq 0} u_n$ une série convergente. La quantité $\lim_{N \rightarrow +\infty} S_N$ est notée $\sum_{n=0}^{+\infty} u_n$ et est appelée **somme de la série**.

Définition 46.4 Reste d'ordre N d'une série convergente

Soit $\sum_{n \geq 0} u_n$ une série de $\mathbb{K}$ ($= \mathbb{R}$ ou $\mathbb{C}$), convergente, de somme S . Soit N un entier naturel. Nous appelons **reste d'ordre N** de cette série la quantité notée R_N et égale à $R_N = S - S_N$. Cette quantité est notée $\sum_{n=N+1}^{+\infty} u_n$.



- Par définition de la convergence de la série, nous avons $\lim_{N \rightarrow +\infty} R_N = 0$.
- L'unicité de la limite implique l'unicité de la somme d'une série convergente.
- Déterminer la **nature** d'une série, c'est dire si elle converge ou si elle diverge. C'est un problème différent que de calculer sa somme en cas de convergence. Parfois les deux problèmes peuvent être traités simultanément. Il est fréquent de prouver qu'une série est convergente sans pouvoir en calculer la somme.

Nous donnons ci-dessous quelques exemples de séries numériques.

1. **Série géométrique** : la série $\sum_{n \geq 0} q^n$, où q appartient à $\mathbb{C}$, est convergente si et seulement si $|q| < 1$. Lorsqu'elle existe, sa somme est égale à $\frac{1}{1-q}$.
2. **Série harmonique** : la série $\sum_{n \geq 1} \frac{1}{n}$ est divergente et $\sum_{k=1}^n \frac{1}{k} \sim \ln(n)$ lorsque $n \rightarrow +\infty$.
3. **Série harmonique alternée** : la série $\sum_{n \geq 1} \frac{(-1)^{n-1}}{n}$ est convergente, de somme $\ln(2)$.
4. **Série des $\frac{1}{n^2}$** : la série $\sum_{n \geq 1} \frac{1}{n^2}$ est convergente, de somme $\frac{\pi^2}{6}$, voir l'Exercice 57.7.
5. **Série exponentielle** : pour tout réel x , la série $\sum_{n \geq 0} \frac{x^n}{n!}$ est convergente, de somme $\exp(x)$.

2. Propriétés des séries convergentes

Proposition 46.1 Convergence des séries à termes complexes

Pour $z \in \mathbb{C}$, nous notons $\Re(z)$ et $\Im(z)$ ses parties réelle et imaginaire. La série complexe $\sum_{n \geq 0} z_n$ est convergente si et seulement si les séries réelles $\sum_{n \geq 0} \Re(z_n)$ et $\sum_{n \geq 0} \Im(z_n)$ le sont. Nous avons alors : $\sum_{n=0}^{+\infty} z_n = \sum_{n=0}^{+\infty} \Re(z_n) + i \sum_{n=0}^{+\infty} \Im(z_n)$.

Proposition 46.2 Une condition nécessaire de convergence

Si la série $\sum_{n \geq 0} u_n$ est convergente, alors $\lim_{n \rightarrow +\infty} u_n = 0$.



- Attention ! La réciproque est fausse. Par exemple, la série $\sum_{n \geq 1} \frac{1}{n}$ diverge mais $\lim_{n \rightarrow +\infty} \frac{1}{n} = 0$.
- Si $\lim_{n \rightarrow +\infty} u_n$ n'existe pas ou est non nulle, la série $\sum_{n \geq 0} u_n$ est grossièrement divergente.

Proposition 46.3 Combinaisons linéaires de séries convergentes

Soient (u_n) et (v_n) deux suites de $\mathbb{K}$ ($= \mathbb{R}$ ou $\mathbb{C}$) et $(\lambda, \mu) \in \mathbb{K}^2$. Si les séries $\sum_{n \geq 0} u_n$ et $\sum_{n \geq 0} v_n$ sont convergentes, alors la série $\sum_{n \geq 0} (\lambda u_n + \mu v_n)$ est convergente et nous avons l'égalité : $\sum_{n=0}^{+\infty} (\lambda u_n + \mu v_n) = \lambda \sum_{n=0}^{+\infty} u_n + \mu \sum_{n=0}^{+\infty} v_n$.



- Si $\lambda \neq 0$, les séries $\sum_{n \geq 0} (\lambda u_n)$ et $\sum_{n \geq 0} u_n$ sont de même nature.
- Si $\sum_{n \geq 0} u_n$ et $\sum_{n \geq 0} v_n$ sont de natures différentes, alors $\sum_{n \geq 0} (u_n + v_n)$ est divergente.
- Il est possible que $\sum_{n \geq 0} (u_n + v_n)$ soit convergente alors que ni $\sum_{n \geq 0} u_n$ ni $\sum_{n \geq 0} v_n$ ne le sont. Nous ne développerons donc pas $\sum_{n \geq 0} (u_n + v_n)$ sans vérifier la convergence de $\sum_{n \geq 0} u_n$ et $\sum_{n \geq 0} v_n$.

Proposition 46.4 Critère de Cauchy pour la convergence d'une série

Soit (u_n) une suite de $\mathbb{K}$ ($= \mathbb{R}$ ou $\mathbb{C}$). La série $\sum_{n \geq 0} u_n$ est convergente si et seulement si :

$$\forall \varepsilon > 0, \exists N \in \mathbb{N} \text{ tel que : } \forall m \geq N, \forall n \geq N, \left| \sum_{k=m}^n u_k \right| \leq \varepsilon.$$

Ce critère peut encore s'écrire :

$$\forall \varepsilon > 0, \exists N \in \mathbb{N} \text{ tel que : } \forall n \geq N, \forall p \geq 0, \left| \sum_{k=n}^{n+p} u_k \right| \leq \varepsilon.$$

3. Convergence absolue et semi-convergence

Définition 46.5 Convergence absolue

Soit (u_n) une suite de $\mathbb{K}$ ($= \mathbb{R}$ ou $\mathbb{C}$). La série $\sum_{n \geq 0} u_n$ est dite **absolument convergente** si la série $\sum_{n \geq 0} |u_n|$ est convergente.

Définition 46.6 Semi-convergence

Toute série convergente sans être absolument convergente est dite **semi-convergente**.

Proposition 46.5

Si $\sum_{n \geq 0} u_n$ est absolument convergente, alors la série $\sum_{n \geq 0} u_n$ est convergente, et l'inégalité suivante a lieu : $\left| \sum_{n=0}^{+\infty} u_n \right| \leq \sum_{n=0}^{+\infty} |u_n|$.



La réciproque de cette proposition est fausse ! Par exemple, $\sum_{n=1}^{+\infty} \frac{(-1)^n}{n}$ converge mais $\sum_{n \geq 1} \frac{1}{n}$ diverge.

Proposition 46.6 Produit de deux séries absolument convergentes

Soient (u_n) et (v_n) deux suites de $\mathbb{K}$ ($= \mathbb{R}$ ou $\mathbb{C}$). Nous définissons la suite (w_n) de $\mathbb{K}$ par : $\forall n \in \mathbb{N}, w_n = \sum_{k=0}^n u_k \times v_{n-k}$. Nous disons que la série $\sum_{n \geq 0} w_n$ est le produit de Cauchy des séries $\sum_{n \geq 0} u_n$ et $\sum_{n \geq 0} v_n$. Si les deux séries $\sum_{n \geq 0} u_n$ et $\sum_{n \geq 0} v_n$ sont absolument convergentes, alors la série $\sum_{n \geq 0} w_n$ est absolument convergente et :

$$\sum_{n=0}^{+\infty} w_n = \sum_{n=0}^{+\infty} u_n \sum_{n=0}^{+\infty} v_n.$$

Théorème 46.1 Critère d'Abel

Si la série $\sum_{n \geq 0} u_n v_n$ est telle que la suite $(u_n)_{n \in \mathbb{N}}$ tend vers 0 en décroissant et qu'il existe un réel $M > 0$ tel que, $\forall n \in \mathbb{N}, \left| \sum_{k=0}^n v_k \right| \leq M$, alors $\sum_{n \geq 0} u_n v_n$ converge.

4. Séries alternées

Définition 46.7

Soit (u_n) une suite de $\mathbb{R}$. Nous disons que la série $\sum_{n \geq 0} u_n$ à termes réels est **alternée** si le signe de $(-1)^n u_n$ est constant.

Théorème 46.2 Théorème spécial des séries alternées

Si la série $\sum_{n \geq 0} u_n$ est une série alternée, et si la suite de terme général $|u_n|$ tend vers 0 en décroissant, alors la série $\sum_{n \geq 0} u_n$ est convergente.



- Sous les hypothèses du Théorème 46.2, les sommes partielles (S_{2N}) et (S_{2N+1}) sont des suites adjacentes et convergent vers la somme de la série.
- Le reste $R_N = \sum_{n=N+1}^{+\infty} u_n$ possède le signe de u_{N+1} et nous avons l'inégalité $|R_N| \leq |u_{N+1}|$.

Application L'énergie potentielle

L'énergie potentielle d'interaction gravitationnelle $U_G(n)$ de n points matériels équidistants est donnée par :

$$U_G(n) = -\frac{Gm^2}{a} \sum_{p=1}^{n-1} \frac{n-p}{p}$$

où G représente la constante gravitationnelle, m la masse d'un point matériel et a la distance entre deux points.

L'énergie potentielle d'interaction électrostatique entre des charges ponctuelles alignées (alternativement k charges positives et k charges négatives) est définie par :

$$V_c = -\frac{ke}{2\pi\epsilon_0 a} \sum_{p=1}^{+\infty} \frac{(-1)^{p-1}}{p}$$

où e représente la charge élémentaire et ϵ_0 la permittivité du vide.

Étudier la convergence des deux séries.

1. Étude de $U_G(n)$. Posons $u_n = \sum_{p=1}^{n-1} \frac{n-p}{p}$. Cherchons la limite de u_n quand n tend vers plus l'infini.

$$u_n = \sum_{p=1}^{n-1} \left(\frac{n}{p} - 1 \right) = n \sum_{p=1}^{n-1} \frac{1}{p} - (n-1) \approx n \ln(n) - (n-1) \quad \text{quand } n \rightarrow +\infty.$$

Donc $U_G(n)$ diverge quand n tend vers plus l'infini.

2. Étude de V_c . D'après le cours, nous avons : $\sum_{p \geq 1} \frac{(-1)^{p-1}}{p} = \ln(2)$. Ainsi V_c converge et :

$$V_c = -\frac{ke}{2\pi\epsilon_0 a} \ln(2).$$

Pour s'entraîner (solutions p. 527)

46.1 Déterminer la nature de la série de terme général :

1. $\frac{(-1)^n}{n \sqrt{n^n}}$;

2. $\frac{(-1)^n}{\ln(n!)} \sum_{k=1}^n \frac{1}{k}$.

46.2

L'objectif de l'exercice est de calculer la somme de la série $\sum_{n=0}^{+\infty} \frac{1}{(4n+1)(4n+3)}$.

1. Montrer, sans essayer de calculer sa somme, que la série $\sum_{n=0}^{+\infty} \frac{1}{(4n+1)(4n+3)}$ converge.

2. Soit N un entier naturel non nul.

- a. Montrer qu'il existe un unique couple $(a, b) \in \mathbb{R}^2$ tel que $\frac{1}{(4X+1)(4X+3)} = \frac{a}{4X+1} + \frac{b}{4X+3}$.

- b. En déduire que : $\sum_{n=0}^N \frac{1}{(4n+1)(4n+3)} = \frac{1}{2} \sum_{n=0}^N \frac{1}{4n+1} - \frac{1}{2} \sum_{n=0}^N \frac{1}{4n+3}$.

3. Soit N un entier naturel non nul.

- a. Calculer, pour $n \in N$, la valeur de l'intégrale $\int_0^1 x^n dx$.

b. En déduire que : $\sum_{n=0}^N \frac{1}{4n+1} = \int_0^1 \sum_{n=0}^N t^{4n} dt$, puis que $\sum_{n=0}^N \frac{1}{4n+1} = \int_0^1 \frac{1-t^{4(N+1)}}{1-t^4} dt$.

c. Montrer que : $\sum_{n=0}^N \frac{1}{(4n+1)(4n+3)} = \int_0^1 \frac{1}{2(1+t^2)} dt - \int_0^1 \frac{t^{4(N+1)}}{2(1+t^2)} dt$.

4. Soit N un entier naturel non nul.

- a. Établir l'encadrement suivant : $0 \leq \int_0^1 \frac{t^{4(N+1)}}{2(1+t^2)} dt \leq \frac{1}{8N+10}$.

b. Montrer que : $\int_0^1 \frac{1}{1+t^2} dt = \frac{\pi}{4}$.

- c. En déduire que la somme de la série $\sum_{n=0}^{+\infty} \frac{1}{(4n+1)(4n+3)}$ est égale à $\frac{\pi}{8}$.

Séries à termes positifs

Proposition 47.1 Convergence par majoration des sommes partielles

Soit (u_n) une suite de $\mathbb{R}_+$. La série $\sum_{n \geq 0} u_n$ est convergente si et seulement si la suite (S_N) de ses sommes partielles, qui est croissante, est majorée. En cas de convergence, nous avons l'égalité :

$$\sum_{n=0}^{+\infty} u_n = \sup_{N \geq 0} (S_N) = \lim_{N \rightarrow +\infty} S_N.$$

Proposition 47.2 Convergence par domination

Soient (u_n) et (v_n) deux suites de $\mathbb{R}_+$. Nous supposons que : $\forall n \in \mathbb{N}, u_n \leq v_n$.

- Si la série $\sum_{n \geq 0} u_n$ diverge, alors la série $\sum_{n \geq 0} v_n$ diverge.
- Si la série $\sum_{n \geq 0} v_n$ converge, alors la série $\sum_{n \geq 0} u_n$ converge. Dans ce cas, nous avons alors l'inégalité suivante :

$$\forall N \in \mathbb{N}, \quad \sum_{n=N}^{+\infty} u_n \leq \sum_{n=N}^{+\infty} v_n.$$

Proposition 47.3 Convergence par équivalence ou par prépondérance

Soient (u_n) et (v_n) deux suites de $\mathbb{R}_+$.

- Si $u_n = o(v_n)$ et si la série $\sum_{n \geq 0} v_n$ converge, alors la série $\sum_{n \geq 0} u_n$ converge.
- Si $u_n \sim v_n$, alors les séries $\sum_{n \geq 0} u_n$ et $\sum_{n \geq 0} v_n$ sont de même nature.



Dans la proposition « $u_n \sim v_n \Rightarrow \sum_{n \geq 0} u_n$ et $\sum_{n \geq 0} v_n$ sont de même nature », il est essentiel que les termes généraux u_n et v_n gardent un signe constant quand $n \rightarrow +\infty$.

Proposition 47.4 Somme des relations de comparaison

Soient (u_n) et (v_n) deux suites de $\mathbb{R}_+$. Notons S_N et S'_N les sommes partielles des séries $\sum_{n \geq 0} u_n$ et $\sum_{n \geq 0} v_n$. De même notons R_N et R'_N les restes d'ordre N . En cas de convergence des séries $\sum_{n \geq 0} u_n$ et $\sum_{n \geq 0} v_n$, nous avons :

- si $u_n = o(v_n)$ alors $R_N = o(R'_N)$.
- si $u_n \sim v_n$ alors $R_N \sim R'_N$.

En cas de divergence des séries $\sum_{n \geq 0} u_n$ et $\sum_{n \geq 0} v_n$, nous avons :

- si $u_n = o(v_n)$ alors $S_N = o(S'_N)$.
- si $u_n \sim v_n$ alors $S_N \sim S'_N$.

Proposition 47.5 Séries de Riemann et séries de Bertrand

- Séries de Riemann : la série $\sum_{n \geq 1} \frac{1}{n^\alpha}$ est convergente si et seulement si $\alpha > 1$.
- Séries de Bertrand : la série $\sum_{n \geq 1} \frac{1}{n^\alpha \ln^\beta n}$ est convergente si et seulement si $\alpha > 1$ ou si $\alpha = 1$ et $\beta > 1$.

Proposition 47.6 Règle de d'Alembert

Soit $\sum_{n \geq 0} u_n$ une série à termes strictement positifs. Nous supposons que $\lim_{n \rightarrow +\infty} \frac{u_{n+1}}{u_n} = \alpha$.

- Si $0 \leq \alpha < 1$, la série $\sum_{n \geq 0} u_n$ est convergente.
- Si $\alpha > 1$, la série $\sum_{n \geq 0} u_n$ est divergente.
- Si $\alpha = 1$, nous ne pouvons rien conclure. C'est le cas litigieux de la règle de d'Alembert. Toutefois, si $\lim_{n \rightarrow +\infty} \frac{u_{n+1}}{u_n} = 1^+$, alors la série $\sum_{n \geq 0} u_n$ est divergente.



La règle de d'Alembert est bien adaptée aux cas où le terme général s'exprime à l'aide de produits, en particulier lorsqu'il contient des puissances ou des factorielles.

Proposition 47.7 Règle de Cauchy

Soit $\sum_{n \geq 0} u_n$ une série à termes strictement positifs. Nous supposons que $\lim_{n \rightarrow +\infty} \sqrt[n]{u_n} = l$ avec $l \in [0; +\infty[$.

- Si $l < 1$, la série $\sum_{n \geq 0} u_n$ est convergente.
- Si $l > 1$, la série $\sum_{n \geq 0} u_n$ est divergente.
- Si $l = 1$, nous ne pouvons rien conclure. C'est le cas litigieux de la règle de Cauchy.



La règle de Cauchy n'est bien adaptée qu'à l'étude des séries dont le terme général contient essentiellement des puissances.

Proposition 47.8 Comparaison d'une série à une intégrale

Soit $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ une fonction continue par morceaux, décroissante. La série $\sum_{n \geq 0} f(n)$ converge si et seulement si la fonction f est intégrable sur $[0; +\infty[$, et, dans ces conditions, nous avons :

$$\forall n \in \mathbb{N}, \quad \int_{n+1}^{+\infty} f(t) dt \leq \sum_{k=n+1}^{+\infty} f(k) \leq \int_n^{+\infty} f(t) dt.$$



Fiche 53

Application**1. Calcul d'une somme d'une série numérique**

Calculer la somme de la série :

$$\sum_{n=1}^{+\infty} \frac{1}{n(n+1)(n+2)}.$$

2. Application directe de la règle de d'Alembert

Étudier la nature de la série numérique $\sum_{n \geq 1} u_n$ de terme général $u_n = \frac{5^n n!}{n^n}$.

3. Application directe de la règle de Cauchy

Étudier la nature de la série numérique $\sum_{n \geq 1} u_n$ de terme général $u_n = \left(\frac{n-1}{3n+2} \right)^{3n}$.

1. Soit $F \in \mathbb{R}(X)$ la fraction rationnelle : $\frac{1}{X(X+1)(X+2)}$. Décomposons F en éléments simples :

$$F(X) = \frac{a}{X} + \frac{b}{X+1} + \frac{c}{X+2} = \frac{1}{2} \frac{1}{X} - \frac{1}{X+1} + \frac{1}{2} \frac{1}{X+2}.$$

Ainsi la somme partielle de la série s'écrit :

$$\begin{aligned} \sum_{k=1}^n F(k) &= \sum_{k=1}^n \left(\frac{1}{2} \frac{1}{k} - \frac{1}{2} \frac{1}{k+1} - \frac{1}{2} \frac{1}{k+1} + \frac{1}{2} \frac{1}{k+2} \right) \\ &= \frac{1}{2} \sum_{k=1}^n \left(\frac{1}{k} - \frac{1}{k+1} \right) + \frac{1}{2} \sum_{k=1}^n \left(\frac{1}{k+2} - \frac{1}{k+1} \right) \\ &= \frac{1}{2} \left(1 - \frac{1}{n+1} \right) + \frac{1}{2} \left(\frac{1}{n+2} - \frac{1}{2} \right) \\ &= \frac{1}{4} - \frac{1}{2} \frac{1}{(n+1)(n+2)}. \end{aligned}$$

Nous en déduisons que

$$\lim_{n \rightarrow +\infty} \sum_{k=1}^n F(k) = \frac{1}{4}$$

et que la somme de la série vaut alors $\sum_{n=1}^{+\infty} \frac{1}{n(n+1)(n+2)} = \frac{1}{4}$.

2.



Pour étudier la nature de la série définie ci-dessus, nous allons utiliser la règle de d'Alembert.

Calculons le rapport $\frac{u_{n+1}}{u_n}$.

$$\frac{u_{n+1}}{u_n} = \frac{5^{n+1}(n+1)!n^n}{(n+1)^{n+1}5^n n!} = 5 \left(\frac{n}{n+1} \right)^n = 5e^{n \ln(1 - \frac{1}{n+1})}.$$

Or nous avons $\ln \left(1 - \frac{1}{n+1} \right) \sim -\frac{1}{n+1}$ au voisinage de $+\infty$. Ainsi

$$\frac{u_{n+1}}{u_n} \sim 5e^{-\frac{n}{n+1}} \quad \text{au voisinage de } +\infty.$$

Ainsi, nous pouvons conclure que :

$$\lim_{n \rightarrow +\infty} \frac{u_{n+1}}{u_n} = \frac{5}{e} > 1.$$

Donc la série $\sum_{n \geq 1} u_n$ diverge.

3.



Pour étudier la nature de la série définie ci-dessus, nous allons utiliser la règle de Cauchy.

$$\lim_{n \rightarrow +\infty} \sqrt[n]{u_n} = \lim_{n \rightarrow +\infty} \sqrt[n]{\left(\frac{n-1}{3n+2} \right)^{3n}} = \lim_{n \rightarrow +\infty} \left(\frac{1 - \frac{1}{n}}{3 + \frac{2}{n}} \right)^3 = \frac{1}{27} < 1.$$

Donc la série $\sum_{n \geq 1} u_n$ converge.

Pour s'entraîner (solutions p. 527)

47.1 Déterminer la nature de la série de terme général :

1. $\frac{3^{2^n}}{2^{3^n}}$;
2. $\ln\left(\frac{n^2 + n + 1}{n^2 + n - 1}\right)$;
3. $(\sqrt{n} + \sqrt[3]{n})^{-\sqrt{n}}$;
4. $\frac{1}{n + (-1)^n \sqrt{n}}$;
5. $\frac{(-1)^n \sqrt{n} + \sin n}{n^2}$;
6. $\left(1 - \frac{n}{\ln n}\right)^{-n}$.

47.2 En utilisant une comparaison d'une série à une intégrale, montrer que la somme partielle de la série harmonique, $\sum_{k=1}^n \frac{1}{k}$, est équivalente à $\ln n$.

47.3 Pour tout $n \in \mathbb{N}^*$, nous notons $u_n = \left(\sum_{k=1}^n \frac{1}{k}\right) - \ln n$. Pour tout $n \in \mathbb{N}^*$, nous notons $v_n = u_{n+1} - u_n$.

1. Montrer que v_n est équivalent, quand n tend vers $+\infty$, à $-\frac{1}{2n^2}$. Indication : faire un développement limité de $\ln\left(1 - \frac{1}{n+1}\right)$.

2. Quelle est la nature de la série $\sum_{n \geq 1} v_n$?

3. En déduire que la suite $(u_n)_{n \geq 1}$ converge. Nous notons γ la limite de la suite $(u_n)_{n \geq 1}$.

En déduire que : $\sum_{k=1}^n \frac{1}{k} = \ln n + \gamma + o(1)$.

47.4 Règle de Raabe et Duhamel

Soit une série $\sum_n u_n$ à termes dans $\mathbb{R}_+^*$. Nous supposons qu'il existe $\alpha \in \mathbb{R}$ tel que :

$$\frac{u_{n+1}}{u_n} = 1 - \frac{\alpha}{n} + o\left(\frac{1}{n}\right) \quad \text{au voisinage de } +\infty.$$

Montrer :

1. si $\alpha > 1$, alors $\sum_n u_n$ converge.

2. si $\alpha < 1$, alors $\sum_n u_n$ diverge.

Soient (f_n) une suite de fonctions définies sur un intervalle I de $\mathbb{R}$ et à valeurs dans $\mathbb{R}$ ou $\mathbb{C}$ et f une fonction définie sur cet intervalle I et à valeurs dans $\mathbb{R}$ ou $\mathbb{C}$.

1. Convergence simple et convergence uniforme

Définition 48.1

La suite de fonctions (f_n) **converge simplement vers une fonction f sur l'intervalle I** si et seulement si, pour tout x de I , la suite $(f_n(x))_{n \in \mathbb{N}}$ converge vers $f(x)$ dans $\mathbb{R}$ ou $\mathbb{C}$, ce qui s'écrit aussi :

$$\forall x \in I, \quad \lim_{n \rightarrow +\infty} f_n(x) = f(x),$$

ou encore :

$$\forall \varepsilon > 0, \forall x \in I, \exists N(\varepsilon, x) \in \mathbb{N}^*, n > N(\varepsilon, x) \implies |f_n(x) - f(x)| < \varepsilon.$$

Nous disons que f est la **limite simple** de $(f_n)_{n \in \mathbb{N}}$.

Exemple

La suite de fonctions définie pour tout $n \in \mathbb{N}^*$ par $(f_n(x) = 3e^{-nx^2})$ converge simplement vers la fonction f sur $\mathbb{R}_+$ définie par :

$$\begin{cases} f(x) = 0 & \text{pour } x > 0 \\ f(0) = 3. \end{cases}$$

En effet, nous avons :

$$\text{pour tout } x > 0, \lim_{n \rightarrow +\infty} 3e^{-nx^2} = 0 \quad \text{et} \quad \text{pour tout } n \in \mathbb{N}^*, f_n(0) = 3.$$

Définition 48.2

f étant la limite simple de la suite des fonctions (f_n) , nous disons que la **convergence de (f_n) vers f est uniforme sur l'intervalle I** si et seulement si

$$\lim_{n \rightarrow +\infty} \|f_n - f\|_\infty = 0, \quad \text{où } \|f_n - f\|_\infty = \sup_{x \in I} |f_n(x) - f(x)|$$

ou encore

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall n \in \mathbb{N}, \forall x \in I, (n \geq N \implies |f_n(x) - f(x)| < \varepsilon).$$

Nous disons que f est la **limite uniforme** de $(f_n)_{n \in \mathbb{N}}$.



Dans la définition de la convergence uniforme, l'entier N ne doit pas dépendre de x .

Exemple

Soient la suite de fonctions définie par $\left(f_n(x) = \frac{\sin(nx)}{n}\right)$ et la fonction f nulle sur $[0; \pi]$.

$$\|f_n - f\|_\infty = \sup_{x \in [0; \pi]} |f_n(x) - f(x)| = \sup_{x \in [0; \pi]} \left| \frac{\sin(nx)}{n} \right| = \frac{1}{n} \xrightarrow{n \rightarrow +\infty} 0.$$

Ainsi, la suite de fonctions (f_n) converge uniformément vers la fonction f nulle sur $[0; \pi]$.



- Le nombre $\|f_n - f\|_\infty$ se calcule souvent avec l'étude des variations de la fonction $f_n - f$. Quand ce calcul est trop difficile, nous chercherons à minorer ou à majorer cette quantité.
- La convergence uniforme entraîne la convergence simple. **Attention, la réciproque est fausse.** Une suite de fonctions peut converger simplement sans converger uniformément.

Exemple

Soit la suite de fonctions $(f_n(x) = x^n)$ avec $x \in [0; 1]$. Nous avons :

$$\text{pour } 0 \leq x < 1, \lim_{n \rightarrow +\infty} x^n = 0 \quad \text{et} \quad \text{pour } x = 1 \lim_{n \rightarrow +\infty} x^n = 1.$$

Ainsi la suite de fonctions (f_n) converge simplement vers la fonction f telle que :

$$\begin{cases} f(x) = 0 & \text{pour } 0 \leq x < 1 \\ f(1) = 1. \end{cases}$$

Or pour tout $n \in \mathbb{N}$, nous avons $\sup_{x \in [0; 1]} |f_n(x) - f(x)| = 1 \neq 0$. La suite de fonctions (f_n) ne converge pas uniformément vers f .

2. Propriétés

Proposition 48.1 Continuité de la limite

Si la suite des fonctions (f_n) converge uniformément vers la fonction f sur l'intervalle I , et si chaque fonction f_n est continue sur l'intervalle I , alors la fonction f est continue sur l'intervalle I .



- Si les fonctions f_n sont continues sur l'intervalle I , et si la fonction limite f n'est pas continue sur I , alors la convergence n'est pas uniforme.
- Il suffit que la convergence soit uniforme sur tout segment inclus dans l'intervalle I , pour que la fonction f soit continue sur l'intervalle I .

Proposition 48.2 Intégration de la limite

Si la suite de fonctions (f_n) converge uniformément vers la fonction f sur l'intervalle I , et si chaque fonction f_n est continue sur l'intervalle I , alors pour tous a et b dans l'intervalle I , nous avons l'égalité suivante :

$$\int_a^b \lim_{n \rightarrow +\infty} f_n(t) dt = \int_a^b f(t) dt = \lim_{n \rightarrow +\infty} \int_a^b f_n(t) dt.$$

Proposition 48.3 Dérivation de la limite

Soit (f_n) une suite de fonctions de classe $\mathcal{C}^1$ sur l'intervalle I , convergeant en un point $a \in I$. Si la suite des fonctions dérivées (f'_n) converge uniformément sur l'intervalle I vers une fonction g , alors la suite des fonctions (f_n) converge uniformément vers une fonction f de classe $\mathcal{C}^1$ sur l'intervalle I qui vérifie :

$$f' = g = \lim f'_n.$$

Application

1. La convergence simple n'implique pas la convergence uniforme

Soit la suite de fonctions (f_n) définies sur $[0; 1]$ par : $f_n(x) = \frac{nx}{1 + nx}$.

a. Montrer que la suite de fonctions (f_n) converge simplement vers la fonction f telle que :

$$\begin{cases} f(x) = 1 & \text{pour } x \in]0; 1] \\ f(0) = 0 \end{cases}$$

b. Montrer que la convergence de la suite de fonctions (f_n) n'est pas uniforme.

2. Encore une fois !

Soit la suite de fonctions définies sur $[0; 1]$ par $\left(f_n(x) = \frac{nx}{1 + n^2 x^4}\right)$. Montrer que la suite de fonctions (f_n) ne converge pas uniformément vers la fonction nulle sur $[0; 1]$.

3. Et un dernier...

Étudier la convergence simple et la convergence uniforme de la suite de fonctions (f_n) définie, pour $x \in \mathbb{R}_+^*$, par $(f_n(x) = nx \exp(-nx^2))$.

1. a. Nous avons d'une part pour tout $x \in]0; 1]$ $\lim_{n \rightarrow +\infty} f_n(x) = \lim_{n \rightarrow +\infty} \frac{nx}{1 + nx} = 1$ et d'autre part pour tout entier n , $f_n(0) = 0$. Donc la suite de fonctions (f_n) converge simplement vers la fonction f donnée dans l'énoncé.

b. Il ne peut pas y avoir convergence uniforme car la limite simple f de la suite de fonctions (f_n) n'est pas continue sur $[0; 1]$ alors que toutes les fonctions f_n le sont.

2. Nous avons d'une part pour tout $x \in [0; 1]$, $\lim_{n \rightarrow +\infty} f_n(x) = \lim_{n \rightarrow +\infty} \frac{x}{\frac{1}{n} + nx^4} = 0$. Donc

la fonction nulle est la limite simple de la suite de fonctions (f_n) . Maintenant déterminons $\sup_{x \in [0; 1]} |f_n(x) - f(x)| = \sup_{x \in [0; 1]} f_n(x)$. Nous calculons la valeur de x pour laquelle $f_n(x)$ est maximale. Pour cela dérivons f_n qui est une fonction dérivable sur $[0; 1]$ étant donné que le dénominateur de f_n ne s'annule jamais. Donc pour tout $x \in [0; 1]$, nous avons $f'_n(x) = \frac{n - 3n^3 x^4}{(1 + n^2 x^4)^2}$. Ensuite, cherchons les points où f'_n s'annule.

$$f'_n(x_n) = 0 \iff x_n = (3n^2)^{-\frac{1}{4}} \quad \text{et} \quad f_n(x_n) = \frac{3^{\frac{3}{4}} \sqrt{n}}{4}.$$

Ainsi nous avons :

$$\sup_{x \in [0; 1]} |f_n(x) - f(x)| = \frac{3^{\frac{3}{4}} \sqrt{n}}{4}$$

qui ne tend pas vers 0 quand n tend vers l'infini, donc la suite de fonctions (f_n) ne converge pas uniformément.

3. Calculons pour $x > 0$ la limite quand n tend vers l'infini de $f_n(x)$. Nous avons $\lim_{n \rightarrow +\infty} f_n(x) = \lim_{n \rightarrow +\infty} nx \exp(-nx^2) = 0$. La suite de fonctions (f_n) converge simplement vers la fonction nulle sur $\mathbb{R}_+^*$.

Pour la convergence uniforme, cherchons $\sup_{x \in]0; +\infty[} |f_n(x)|$. La fonction f_n étant dérivable sur $\mathbb{R}_+$, nous avons :

$$f'_n(x) = n \exp(-nx^2) (1 - 2nx^2).$$

$f'_n(x)$ s'annule pour $x_n = \frac{1}{\sqrt{2n}}$. La dérivée est positive sur $[0; x_n]$ et négative sur $[x_n; +\infty[$.

Donc, f_n est strictement croissante sur $[0; x_n]$ et strictement décroissante sur $[x_n; +\infty[$. La fonction f_n admet un maximum en $x_n > 0$. Nous en déduisons :

$$\sup_{x \in]0; +\infty[} |f_n(x) - 0| = \sup_{x \in]0; +\infty[} f_n(x) = f_n(x_n) = \sqrt{\frac{n}{2}} \exp\left(-\frac{1}{2}\right).$$

Comme $\lim_{n \rightarrow +\infty} f_n(x_n) = +\infty$, la suite de fonctions (f_n) ne converge pas uniformément sur $\mathbb{R}_+^*$.

Pour s'entraîner (solutions p. 528)

48.1 Soit la suite de fonctions

$\left(f_n(x) = \frac{ne^{-x} + 3x^3}{2n + x}\right)$, pour tout $x \in [0; 1]$ et pour tout $n \in \mathbb{N}^*$. Étudier la convergence simple et la convergence uniforme de la suite de fonctions (f_n) sur $[0; 1]$.

48.2 

1. Montrer que la suite de fonctions

$\left(f_n(x) = \frac{x}{1 + nx}\right)$ pour $x \in [0; 1]$ converge simplement et uniformément vers la fonction nulle sur $[0; 1]$.

2. Montrer que

$$\lim_{n \rightarrow +\infty} \int_0^1 f_n(x) dx = \int_0^1 \lim_{n \rightarrow +\infty} f_n(x) dx.$$

Séries de fonctions

Soit (f_n) une suite de fonctions définies sur un intervalle I de $\mathbb{R}$ et à valeurs dans $\mathbb{R}$ ou $\mathbb{C}$. Nous considérons les sommes partielles définies par : $S_n(x) = \sum_{k=0}^n f_k(x)$ pour $x \in I$.

1. Convergence simple, convergence uniforme, convergence normale

Définition 49.1

La **série de fonctions** $\sum_n f_n$ **converge simplement sur l'intervalle** I si et seulement si la suite (S_n) des sommes partielles converge simplement sur l'intervalle I , c'est-à-dire si et seulement si, pour chaque x de I , la série $\sum_n f_n(x)$ converge dans $\mathbb{R}$ ou $\mathbb{C}$, ce qui s'écrit aussi :

$$S(x) = \lim_{n \rightarrow +\infty} S_n(x) = \sum_{k=0}^{+\infty} f_k(x), \quad \text{pour tout } x \in I.$$

Définition 49.2

Soit $\sum_n f_n$ une série de fonctions qui converge simplement sur I .

- Pour chaque n de $\mathbb{N}$, le **reste d'ordre** n est l'application $R_n : I \rightarrow \mathbb{R}$ ou $\mathbb{C}$ définie par :

$$\forall x \in I, \quad R_n(x) = \sum_{k=n+1}^{+\infty} f_k(x).$$

- La **somme de la série** $\sum_n f_n$ est l'application, notée $\sum_{n=0}^{+\infty} f_n$, définie de I dans $\mathbb{R}$ ou $\mathbb{C}$ par :

$$\forall x \in I, \quad \left(\sum_{n=0}^{+\infty} f_n \right)(x) = \sum_{n=0}^{+\infty} f_n(x).$$

Définition 49.3

La **série de fonctions** $\sum_n f_n$ **converge absolument sur l'intervalle** I si et seulement si, pour chaque x de I , la série $\sum_n |f_n(x)|$ converge dans $\mathbb{R}$.

Définition 49.4

La **série de fonctions** $\sum_n f_n$ **converge uniformément sur l'intervalle** I si et seulement si la suite (S_n) des sommes partielles converge uniformément sur l'intervalle I .

Proposition 49.1

La série de fonctions $\sum_n f_n$ converge uniformément sur l'intervalle I si et seulement si la série $\sum_n f_n$ converge simplement sur I et la suite des restes (R_n) converge uniformément vers 0 sur I .

Proposition 49.2

Si la série de fonctions $\sum_n f_n$ converge uniformément sur l'intervalle I , alors la suite (f_n) converge uniformément vers 0 sur I .

Théorème 49.1 CNS de Cauchy de convergence uniforme d'une série de fonctions

Pour que $\sum_n f_n$ converge uniformément sur l'intervalle I , il faut et il suffit que :

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall (p, q) \in \mathbb{N}^2, \forall x \in I, \left(N \leq p < q \Rightarrow \left| \sum_{k=p+1}^q f_k(x) \right| \leq \varepsilon \right).$$

Théorème 49.2 Critère d'Abel

Si la série de fonctions de terme général $(f_n = a_n g_n)$ définie sur l'intervalle I vérifie :

- $\forall n \in \mathbb{N}, a_n \in \mathbb{R}$ et g_n est une fonction bornée sur I ;
- (a_n) converge vers 0 en décroissant ;
- la suite des sommes partielles de $\sum_{n \geq 0} g_n$ est bornée : $\exists A > 0, \forall x \in I, \forall n \in \mathbb{N}, \left| \sum_{k=0}^n g_k(x) \right| \leq A$.

Alors la série de fonctions $\sum_{n \in \mathbb{N}} f_n = \sum_{n \in \mathbb{N}} a_n g_n$ converge uniformément sur I .

Définition 49.5

La série de fonctions $\sum_n f_n$ converge normalement sur l'intervalle I si et seulement si la série des normes $\sum_n \|f_n\|_\infty$ converge, où $\|f_n\|_\infty = \sup_{x \in I} |f_n(x)|$.

Proposition 49.3

La série de fonctions $\sum_n f_n$ converge normalement sur l'intervalle I si et seulement s'il existe une série numérique à termes positifs a_n telle que :

$$\forall n \in \mathbb{N}, \quad \forall x \in I, \quad |f_n(x)| \leq a_n \quad \text{et} \quad \sum_{n=0}^{+\infty} a_n \quad \text{convergente.}$$



La recherche de a_n peut se faire par majoration ou en étudiant les variations de f_n .

Théorème 49.3

Si la série de fonctions $\sum_n f_n$ est normalement convergente, alors $\sum_n f_n$ est uniformément convergente et $\sum_n f_n$ est aussi absolument convergente. En particulier, $\sum_n f_n$ est simplement convergente.

2. Propriétés**Théorème 49.4 Continuité de la série**

Soient $a \in I$, $\sum_n f_n$ une série de fonctions définies sur I et à valeurs réelles ou complexes. Si pour tout n de $\mathbb{N}$, f_n est continue en a , (resp. sur I) et si $\sum_n f_n$ converge uniformément sur I alors $\sum_{n=0}^{+\infty} f_n$ est continue en a (resp. sur I).

Théorème 49.5 Intégration sur un segment de la série

Soit $\sum_n f_n$ une série de fonctions continues par morceaux, à valeurs réelles ou complexes et intégrables sur un intervalle I . Supposons que la série converge simplement sur l'intervalle I , de somme $x \mapsto S(x) = \sum_{n=0}^{+\infty} f_n(x)$, continue par morceaux, et que la série $\sum_n \int_I |f_n|$ converge. Alors S est intégrable et nous avons alors l'égalité suivante :

$$\int_I \sum_{n=0}^{+\infty} f_n(x) dx = \int_I S(x) dx = \sum_{n=0}^{+\infty} \int_I f_n(x) dx$$

Théorème 49.6 Dérivation de la série

Soit $\sum_n f_n$ une série de fonctions définies sur I et à valeurs réelles ou complexes. Si pour tout n de $\mathbb{N}$, f_n est de classe $\mathcal{C}^1$ sur I , $\sum_{n \geq 0} f_n$ converge simplement sur I , et $\sum_{n \geq 0} f'_n$ converge uniformément sur tout segment I , alors $\sum_{n \geq 0} f_n$ converge uniformément sur tout segment de I , $\sum_{n=0}^{+\infty} f_n$ est de classe $\mathcal{C}^1$ sur I , et $(\sum_{n=0}^{+\infty} f_n)' = \sum_{n=0}^{+\infty} f'_n$.

Application

1. Convergences simple, normale et uniforme

Soient $\alpha > 0$ et la série de fonctions $\sum_{n \geq 0} f_n$ où f_n est définie sur $[0; +\infty[$ par $f_n(x) = nx^\alpha \exp(-nx^2)$.

a. Étudier, suivant les valeurs de α , la convergence simple et uniforme de la suite de fonctions (f_n) .

b. Étudier, suivant les valeurs de α , la convergence simple de la série de fonctions $\sum_{n \geq 0} f_n$.

c. Étudier, suivant les valeurs de α , la convergence normale de $\sum_{n \geq 0} f_n$.

2. Convergences simple et normale

a. Étudier la convergence simple de la série de fonctions $\sum_{n=0}^{+\infty} \frac{x \cos(x)}{(1+x)^n}$.

b. Étudier la convergence normale de la série de fonctions $\sum_{n=0}^{+\infty} \frac{\sin(3^n x)}{4^n}$. La somme est-elle continue ?

3. Soient $\alpha \in]0; 1[$, $r \in]0; \pi[$ et la série de fonctions complexes de terme général $f_n(x) = \frac{\exp(inx)}{(3n+2)^\alpha}$ pour $x \in I_r = [r; 2\pi - r]$. Montrer que cette série converge uniformément sur I_r .

1. a. Convergences simple et uniforme de la suite

Pour $x \geq 0$ et $\alpha > 0$, $f_n(x)$ tend vers 0 quand n tend vers $+\infty$. Donc, pour tout $\alpha > 0$, la suite de fonctions (f_n) converge simplement sur $\mathbb{R}_+$ vers la fonction nulle. Maintenant étudions la convergence uniforme de (f_n) . Remarquons que $f_0 = 0$ et que pour tout $n \in \mathbb{N}^*$, f_n est strictement positive sur $\mathbb{R}_+^*$ et $f_n(0) = 0$. Ainsi, le maximum de f_n , s'il existe, est atteint sur $\mathbb{R}_+^*$. Or f_n est une fonction de classe $\mathcal{C}^1$ sur $\mathbb{R}_+^*$ et nous avons, pour $n \geq 1$:

$$\forall x \in \mathbb{R}_+^*, \quad f'_n(x) = nx^{\alpha-1} \exp(-nx^2) (\alpha - 2nx^2) \quad \text{d'où} \quad f'_n(x_n) = 0 \iff x_n = \sqrt{\frac{\alpha}{2n}}.$$

Ainsi f_n est strictement croissante sur $]0; \sqrt{\frac{\alpha}{2n}}]$ et strictement décroissante sur $[\sqrt{\frac{\alpha}{2n}}; +\infty[$.

Ainsi f_n admet un maximum en x_n et $\|f_n\|_\infty = f_n(x_n) = \left(\frac{\alpha}{2e}\right)^{\frac{\alpha}{2}} \frac{1}{n^{\frac{\alpha}{2}-1}}$. La suite converge donc uniformément vers 0 si et seulement si $\alpha > 2$.

b. Convergence simple de la série

Remarquons que, comme $x \geq 0$, $\sum_{n \geq 0} f_n(x)$ est une série à termes positifs. De plus, pour $x \geq 0$, $\alpha > 0$ fixés, nous avons $n^2 f_n(x) = n^3 x^\alpha \exp(-nx^2)$ tend vers 0 quand n tend vers $+\infty$, c'est-à-dire, $f_n(x) = O\left(\frac{1}{n^2}\right)$ pour $n \rightarrow +\infty$. Comme la série $\sum_{n \geq 1} \frac{1}{n^2}$ converge, il en est de même de la série $\sum_{n \geq 0} f_n(x)$ (voir Proposition 47.3). Donc, pour tout $\alpha > 0$, $\sum_{n \geq 0} f_n$ converge simplement sur $\mathbb{R}_+$.

c. Convergence normale de la série

D'après la première question, nous avons pour tout $x \in \mathbb{R}_+$, $f_n(x) \leq f_n(x_n)$. La série de terme général $\frac{1}{n^{\frac{\alpha}{2}-1}}$ est une série de Riemann qui converge si et seulement si $\alpha > 4$. Donc, si $\alpha > 4$, la série de fonctions $\sum_{n \geq 0} f_n$ converge normalement, donc uniformément, sur $\mathbb{R}_+$.

2. a. La série de fonctions $\sum_{n=0}^{+\infty} \frac{x \cos(x)}{(1+x)^n}$ converge si et seulement si $x \cos(x)$ est nul ou si la série $\sum_{n=0}^{+\infty} \frac{1}{(1+x)^n}$ converge. Cette dernière est une série géométrique de raison $\frac{1}{1+x}$ qui converge si et seulement si $\left|\frac{1}{1+x}\right| < 1$ donc si et seulement si $x > 0$ ou $x < -2$. D'autre part, $x \cos(x) = 0 \iff x \equiv 0[2\pi]$ ou $x \equiv \frac{\pi}{2}[\pi]$. Donc la série de fonctions converge simplement sur $] -\infty; -2[\cup \left\{-\frac{\pi}{2}\right\} \cup [0; +\infty[$.

b. La série $\sum_{n=0}^{+\infty} \frac{1}{4^n}$ converge car c'est une série géométrique de raison $r = \frac{1}{4} < 1$. De plus, pour tout $x \in \mathbb{R}$, nous avons $\left|\frac{\sin(3^n x)}{4^n}\right| \leq \frac{1}{4^n}$. Donc la série de fonctions $\sum_{n=0}^{+\infty} \frac{\sin(3^n x)}{4^n}$ converge normalement sur $\mathbb{R}$. Ainsi la somme est continue sur $\mathbb{R}$ car elle est la somme uniformément convergente de fonctions continues sur $\mathbb{R}$ (voir Théorème 49.4).

3.



Posons, $\forall x \in I_r$, $f_n(x) = a_n g_n(x)$ avec $a_n = \frac{1}{(3n+2)^n}$ et $g_n(x) = \exp(inx)$ et cherchons à appliquer le critère d'Abel 49.2 pour montrer que la série de fonctions converge uniformément sur I_r .

$\forall n \in \mathbb{N}$, $a_n \in \mathbb{R}$, et $\forall x \in I_r$, $|g_n(x)| = 1$ donc g_n est bornée sur I_r .

De plus, $\forall n \in \mathbb{N}$, $a_n \neq 0$ et $\frac{a_n}{a_{n+1}} = \left(\frac{3n+5}{3n+2}\right)^n > 1$. Ainsi (a_n) est décroissante et $\lim_{n \rightarrow +\infty} a_n = 0$.

$\sum_{i=0}^n g_i(x) = 1 + \exp(ix) + (\exp(ix))^2 + \dots + (\exp(ix))^n = \frac{1 - (\exp(ix))^{n+1}}{1 - \exp(ix)}$, $\exp(ix) \neq 1$ car

$x \in I_r$, et nous avons : $\left|\sum_{i=0}^n g_i(x)\right| = \left|\frac{1 - (\exp(ix))^{n+1}}{1 - \exp(ix)}\right| \leq \frac{1 + |(\exp(ix))^{n+1}|}{|1 - \exp(ix)|} \leq \frac{2}{|1 - \exp(ix)|}$. Enfin,

$|1 - \exp(ix)| = \sqrt{(1 - \cos(x))^2 + \sin^2(x)} = \sqrt{2 - 2 \cos x} = \sqrt{2 \left(1 - \cos\left(\frac{x}{2}\right)^2 + \sin\left(\frac{x}{2}\right)^2\right)} =$

$2 \left|\sin\left(\frac{x}{2}\right)\right| = 2 \sin\left(\frac{x}{2}\right)$ car $0 < \frac{r}{2} \leq \frac{x}{2} \leq \pi - \frac{r}{2} < \pi$. Si $0 < \frac{r}{2} \leq \frac{x}{2} \leq \frac{\pi}{2}$, la fonction sinus est croissante sur $]0; \frac{\pi}{2}]$, nous avons alors $\sin\left(\frac{x}{2}\right) \geq \sin\left(\frac{r}{2}\right)$. Si $\frac{\pi}{2} \leq \frac{x}{2} \leq \pi - \frac{r}{2} < \pi$, la fonction sinus est décroissante sur $[\frac{\pi}{2}; \pi[$, donc nous avons $\sin\left(\frac{x}{2}\right) \geq \sin\left(\frac{r}{2}\right)$. Donc $\left|\sum_{i=0}^n g_i(x)\right| \leq \frac{1}{2 \sin\left(\frac{r}{2}\right)}$.

La série converge donc uniformément sur I_r .

Pour s'entraîner

49.1



1. Étudier la convergence simple sur $\mathbb{R}_+$ de la série de terme général $f_n(x) = nx^n$.

2. Étudier la convergence simple sur $\mathbb{R}_+$ de la

série de terme général $f_n(x) = \frac{\exp(-nx)}{\ln(n)}$, avec $n \geq 2$.

3. Étudier la convergence sur $\mathbb{R}$ de la série de terme général égal à $f_n(x) = \frac{1}{n^2} \arctan(nx)$, avec $n \geq 1$.

1. Définition et convergence d'une série entière

Définition 50.1

Une série entière est une série de terme général $a_n z^n$, où $(a_n)_{n \in \mathbb{N}}$ est une suite de nombres réels ou complexes et z une variable réelle ou complexe.



L'usage veut que nous adoptions la notation $\sum_{n \geq 0}$ pour parler d'une série entière, tandis que nous écrivons $\sum_{n=0}^{+\infty}$ pour son éventuelle somme, en cas de convergence, pour un z donné.

Définition 50.2 Rayon de convergence

Si $\sum_{n \geq 0} a_n z^n$ est une série entière, elle vérifie une et une seule des trois propriétés suivantes :

1. la série entière converge uniquement pour $z = 0$.
2. il existe un nombre réel $R > 0$ tel que la série entière converge absolument pour tout z tel que $|z| < R$, et diverge pour tout z tel que $|z| > R$.
3. la série entière converge absolument pour tout z .

Le **rayon de convergence** de la série entière est le nombre réel R qui apparaît dans le deuxième cas. Dans le premier cas, $R = 0$ et dans le troisième cas, $R = +\infty$.

Pour déterminer le rayon de convergence R , nous pouvons utiliser :

- la définition de R , qui est la borne supérieure des ensembles suivants :

$$R = \sup\{r \in \mathbb{R}_+; \sum_{n \geq 0} a_n z^n \text{ converge}\} = \sup\{r \in \mathbb{R}_+; |a_n| r^n \text{ borné}\}.$$

- ou la **formule de Hadamard** qui donne l'expression du rayon de convergence en termes de limite si celle-ci existe :

$$\lim_{n \rightarrow +\infty} (|a_n|^{1/n}) = \frac{1}{R}.$$

- ou la **règle de d'Alembert**. Si les a_n sont non nuls, alors nous calculons la quantité suivante :

$$\frac{|a_{n+1} z^{n+1}|}{|a_n z^n|} = \frac{|a_{n+1}|}{|a_n|} |z|.$$

Si $\frac{|a_{n+1}|}{|a_n|}$ converge vers $l \in \mathbb{R} \cup \{+\infty\}$, la suite $\left(\frac{|a_{n+1}|}{|a_n|} |z|\right)$ converge vers $l|z|$. D'après la règle de d'Alembert, la série converge si $l|z| < 1$, soit $|z| < \frac{1}{l}$, et diverge si $l|z| > 1$, soit $|z| > \frac{1}{l}$. Nous en déduisons que

$$R = \frac{1}{l} = \lim_{n \rightarrow +\infty} \frac{|a_n|}{|a_{n+1}|}$$

avec comme convention que $\frac{1}{0} = +\infty$ et $\frac{1}{+\infty} = 0$.



Fiche 47

Nous pouvons trouver bien des variantes pour la détermination de R , dans la littérature à ce sujet.

Exemples

1. Étudions la série entière de terme général z^n . Nous savons que :

$$\sum_{n=0}^N z^n = z^0 + z^1 + \cdots + z^n + \cdots + z^N = \frac{1 - z^{N+1}}{1 - z}.$$

Si $|z| < 1$, alors $\lim_{N \rightarrow +\infty} z^{N+1} = 0$. Ainsi nous avons $\sum_{n=0}^{+\infty} z^n = \frac{1}{1-z}$ et donc $R \geq 1$. Remarquons que pour $z = 1$, la série entière diverge. Donc $R = 1$.

2. Étudions la série entière de terme général $n^n z^n$, nous avons $\sqrt[n]{|n^n|} = n \rightarrow +\infty$ si $n \rightarrow +\infty$ et donc $R = 0$.

Proposition 50.1 Convergence absolue

La série de rayon de convergence R converge absolument dans l'intervalle ouvert de convergence $] -R; +R[$ dans le cas réel ; dans le disque ouvert de convergence $\{z \in \mathbb{C}; |z| < R\}$ dans le cas complexe.



- Pour $|z| = R$, il n'y a pas de résultat général.
- Pour $|z| > R$, la série entière diverge.

Proposition 50.2 Convergence normale

Soient $R \in [0; +\infty]$ et r un réel tel que $0 \leq r < R$. La série de rayon de convergence R converge normalement, donc uniformément dans tout intervalle fermé $[r; r]$ dans le cas réel ; dans tout disque fermé $\{z \in \mathbb{C}; |z| \leq r\}$ dans le cas complexe.

2. Opérations sur les séries entières

Soient $\sum_{n \geq 0} a_n z^n$ et $\sum_{n \geq 0} b_n z^n$ deux séries entières de rayons de convergence respectifs, R_1 et R_2 et de fonctions sommes respectives $z \mapsto f(z) = \sum_{n=0}^{+\infty} a_n z^n$ et $z \mapsto g(z) = \sum_{n=0}^{+\infty} b_n z^n$.

Pour tous α et β des réels, la série entière $\sum_{n \geq 0} (\alpha a_n + \beta b_n) z^n$ a pour fonction somme $z \mapsto \alpha f(z) + \beta g(z)$. Son rayon de convergence R est tel que :

$$R = \min(R_1; R_2) \quad \text{si} \quad R_1 \neq R_2$$

$$R \geq R_1 \quad \text{si} \quad R_1 = R_2.$$

Si nous posons $c_n = a_0 b_n + a_1 b_{n-1} + \cdots + a_n b_0 = \sum_{k=0}^n a_k b_{n-k}$ alors la série entière $\sum_{n \geq 0} c_n z^n$ a pour fonction somme $z \mapsto f(z) \times g(z)$. Son rayon de convergence R est tel que $R \geq \min(R_1; R_2)$.

Proposition 50.3 Continuité

Soit $\sum_{n \geq 0} a_n z^n$ une série entière de rayon de convergence $R \neq 0$ et de fonction somme $z \mapsto f(z)$. La fonction f est continue sur son disque de convergence (cas complexe) ou sur son intervalle de convergence (cas réel).

Les deux propositions suivantes se placent dans le cas réel.

Proposition 50.4 Dérivation

Si la série entière $\sum_{n \geq 0} a_n x^n$ a pour rayon de convergence $R \neq 0$ et pour fonction somme $x \mapsto f(x)$, alors la fonction f est dérivable sur l'intervalle $] -R; +R[$ et nous avons :

$$f'(x) = \sum_{n=1}^{+\infty} n a_n x^{n-1}.$$

Il en résulte que la fonction f est indéfiniment dérivable sur l'intervalle $] -R; +R[$.

Proposition 50.5 Intégration

Si la série entière $f(x) = \sum_{n \geq 0} a_n x^n$ a pour rayon de convergence $R \neq 0$ et pour fonction somme $x \mapsto f(x)$, alors pour tout $x \in] -R; +R[$, nous avons :

$$\int_0^x f(t) dt = \sum_{n=0}^{+\infty} a_n \frac{x^{n+1}}{n+1}.$$

La série entière obtenue par intégration terme à terme a le même rayon de convergence que la série initiale.

Application

1. Le champ électrique total E qui règne à l'intérieur d'un condensateur plan constitué par des armatures métalliques circulaires de rayon R peut se mettre sous la forme :

$$E = \sum_{n=0}^{+\infty} (-1)^n \frac{z^{2n}}{(n!)^2} E_0 = \sum_{n=0}^{+\infty} E_n z^n,$$

où z dépend de la vitesse de la lumière et de la pulsation de la tension appliquée au condensateur, et E_0 est le champ initial entre les armatures. Déterminer le rayon de convergence de cette série entière.

2. Déterminer le rayon de convergence des séries entières suivantes $\sum_{n=1}^{+\infty} a_n z^n$:

$$(a) a_n = \frac{3(-1)^n}{(2n+1)^n}; (b) a_n = \frac{(n-1)^n}{(n+1)!}; (c) a_n = \frac{2}{\sqrt{n}}.$$

3. Les fonctions de Bessel sont très utilisées en physique. Soit la fonction de Bessel d'indice 0 :

$$J_0(x) = \sum_{n=0}^{+\infty} \frac{(-1)^n}{2^{2n}(n!)^2} x^{2n}.$$

Déterminer son domaine de définition en calculant son rayon de convergence.

1.



Commençons par remarquer que tous les termes impairs E_{2p+1} sont nuls : nous ne pouvons donc pas appliquer directement la règle de d'Alembert pour les séries entières car le rapport $\frac{E_{n+1}}{E_n}$ n'est pas défini pour tout $n \in \mathbb{N}$.

Considérons alors la série $E = \sum_{n=0}^{+\infty} E_{2n} z^{2n}$ à laquelle nous appliquons la règle de d'Alembert pour les séries :

$$\lim_{n \rightarrow +\infty} \left| \frac{E_{2n+2} z^{2(n+1)}}{E_{2n} z^{2n}} \right| = \lim_{n \rightarrow +\infty} \left| \frac{E_{2n+2}}{E_{2n}} z^2 \right| = \lim_{n \rightarrow +\infty} \frac{z^2}{(n+1)^2} = 0.$$

Ainsi la série converge pour tout z réel et le rayon de convergence de la série entière E est $R = +\infty$.

2. a. Pour $z \neq 0$ et tout $n \in \mathbb{N}$, nous avons $\sqrt[n]{|a_n z^n|} = \frac{3^{\frac{1}{n}}}{2n+1} |z|$. Donc $\lim_{n \rightarrow +\infty} \sqrt[n]{|a_n z^n|} = 0$. La série entière converge pour tout z et donc $R = +\infty$.



Nous avons utilisé la règle de Cauchy. La formule d'Hadamard allait plus vite, mais nous avons préféré montrer ce petit raisonnement.

Fiche 47

b. Nous utilisons la règle de d'Alembert. Pour $z \neq 0$ et tout $n \in \mathbb{N}$, nous avons $a_n \neq 0$ et

$$\left| \frac{a_{n+1} z^{n+1}}{a_n z^n} \right| = \left| \frac{a_{n+1}}{a_n} \right| |z| = \frac{n}{n+2} \left(\frac{n}{n-1} \right)^n |z|.$$

D'autre part, nous avons $\lim_{n \rightarrow +\infty} \left(\frac{n}{n-1} \right)^n = \lim_{n \rightarrow +\infty} \exp \left(-n \ln \left(\frac{n-1}{n} \right) \right) = \exp(1)$, puisque $\ln \left(1 - \frac{1}{n} \right) = -\frac{1}{n} + o \left(\frac{1}{n} \right)$. Ainsi $\lim_{n \rightarrow +\infty} \left| \frac{a_{n+1}}{a_n} \right| = \exp(1) |z|$. La série entière converge pour $\exp(1) |z| < 1$ et diverge pour $\exp(1) |z| > 1$, donc $R = \frac{1}{\exp(1)}$.

c. Pour $z \neq 0$ et tout $n \in \mathbb{N}$, nous avons $a_n \neq 0$ et $\left| \frac{a_{n+1} z^{n+1}}{a_n z^n} \right| = \sqrt{\frac{n}{n+1}} |z|$ donc $\lim_{n \rightarrow +\infty} \left| \frac{a_{n+1} z^{n+1}}{a_n z^n} \right| = |z|$.

La série entière converge si $|z| < 1$ et diverge si $|z| > 1$, donc le rayon de convergence est $R = 1$.

3. Nous posons : $a_n = \frac{(-1)^n}{2^{2n}(n!)^2}$. Nous avons :

$$\left| \frac{a_{n+1}}{a_n} \right| = \frac{1}{4(n+1)^2} \quad \text{et ainsi} \quad \lim_{n \rightarrow +\infty} \frac{1}{4(n+1)^2} = 0 \quad \text{donc} \quad R = +\infty.$$

Le domaine de définition de la fonction de Bessel J_0 est $\mathbb{R}$.

Pour s'entraîner (solutions p. 529)

50.1 Calculer le rayon et l'intervalle de convergence de la série : $\sum_{n=0}^{+\infty} \frac{(-3)^n}{\sqrt{n+1}} x^n$.

50.2 Calculer le rayon de convergence de la série : $\sum_{n=0}^{+\infty} \frac{2^{n^2}}{(2n)!} x^{2n}$.

50.3 Calculer l'intervalle de convergence de la série : $\sum_{n=0}^{+\infty} \frac{n(x+2)^n}{3^{n+1}}$.

50.4 Calculer le rayon et l'intervalle de convergence de la série : $\sum_{n=1}^{+\infty} \frac{\text{sh}(n)}{\text{ch}(n)} x^n$.

Développement d'une fonction en série entière

1. Développement d'une fonction en série entière

Définition 51.1

- Soit f une fonction d'une variable réelle, définie sur un intervalle ouvert I contenant l'origine. Nous disons que la **fonction** f est **développable en série entière (centrée) en 0** (en abrégé : dSE(0)) si et seulement s'il existe une série entière $\sum_{n \geq 0} a_n x^n$ de rayon de convergence $R \neq 0$ telle que :

$$\forall x \in]-R; +R[\subset I \quad f(x) = \sum_{n=0}^{+\infty} a_n x^n.$$

Nous disons également que la **fonction** f est **analytique**.

- Soit f une fonction d'une variable réelle, définie sur un intervalle ouvert I contenant un point x_0 . Nous disons que la **fonction** f est **développable en série entière (centrée) en x_0** (en abrégé : dSE(x_0)) si et seulement s'il existe une série entière $\sum_{n \geq 0} a_n (x - x_0)^n$ de rayon de convergence $R \neq 0$ telle que :

$$\forall x \in]x_0 - R; x_0 + R[\subset I \quad f(x) = \sum_{n=0}^{+\infty} a_n (x - x_0)^n.$$



Si le développement en série entière de la fonction f existe, alors celui-ci est unique.

Proposition 51.1

Si la fonction f est développable en série entière en 0 où $f(x) = \sum_{n=0}^{+\infty} a_n x^n$ pour tout $x \in]-R; +R[$, alors la fonction f est indéfiniment dérivable sur $] - R; +R[$ et $\forall n \in \mathbb{N}, \quad a_n = \frac{f^{(n)}(0)}{n!}$.

Nous pouvons énoncer une réciproque de cette condition.

Proposition 51.2

Si la fonction f est indéfiniment dérivable dans l'intervalle I contenant x_0 et s'il existe une constante M telle que : $\forall n \in \mathbb{N}, \quad \forall x \in I, \quad |f^{(n)}(x)| \leq M$ alors la fonction f est analytique en x_0 .

Définition 51.2

- Soit f une fonction d'une variable réelle, définie sur un intervalle ouvert I contenant l'origine et dSE(0). La relation $f(x) = \sum_{n=0}^{+\infty} \frac{f^{(n)}(0)}{n!} x^n$, qui est valable sur un voisinage de 0, est appelée le **développement en série entière de f en 0** (en abrégé : DSE(0)), ou **développement de Mac-Laurin de f** .

- Soient $x_0 \in \mathbb{R}$ et f une fonction d'une variable réelle, définie sur un intervalle ouvert I contenant x_0 et $\text{DSE}(x_0)$. La relation $f(x) = \sum_{n=0}^{+\infty} \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n$, qui est valable sur un voisinage de x_0 , est appelée le **développement en série entière de f en x_0** (en abrégé : $\text{DSE}(x_0)$), ou **développement de Taylor de f en x_0** .

Proposition 51.3

Soient f une fonction d'une variable réelle, définie sur un intervalle ouvert I contenant l'origine, développable en série entière en 0 et $\sum_{n \geq 0} a_n x^n$ le $\text{DSE}(0)$ de f .

- Si f est paire, alors le $\text{DSE}(0)$ de f est pair, c'est-à-dire : $\forall p \in \mathbb{N}, a_{2p+1} = 0$.
- Si f est impaire, alors le $\text{DSE}(0)$ de f est impair, c'est-à-dire : $\forall p \in \mathbb{N}, a_{2p} = 0$.



De façon plus générale, la fonction f est analytique au point x_0 s'il existe un intervalle ouvert contenant le point x_0 dans lequel la somme de sa série de Taylor au point x_0 est égale à la fonction $f(x)$.

2. Quelques développements de base

Nous donnons les principaux développements en série entière en 0. Certains se déduisent aisément de ceux que nous donnons ici.

$\exp(x) = \sum_{n=0}^{+\infty} \frac{x^n}{n!}$	$R = +\infty$
$\cos(x) = \sum_{n=0}^{+\infty} (-1)^n \frac{x^{2n}}{(2n)!}$	$R = +\infty$
$\sin(x) = \sum_{n=0}^{+\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}$	$R = +\infty$
$\text{ch}(x) = \sum_{n=0}^{+\infty} \frac{x^{2n}}{(2n)!}$	$R = +\infty$
$\text{sh}(x) = \sum_{n=0}^{+\infty} \frac{x^{2n+1}}{(2n+1)!}$	$R = +\infty$
$\frac{1}{1-x} = \sum_{n=0}^{+\infty} x^n$	$R = 1$
$\ln(1+x) = \sum_{n=1}^{+\infty} (-1)^{n-1} \frac{x^n}{n}$	$R = 1$
$\arctan(x) = \sum_{n=0}^{+\infty} (-1)^n \frac{x^{2n+1}}{2n+1}$	$R = 1$
$(1+x)^m = \sum_{n=0}^m C_m^n x^n = \sum_{k=0}^m \binom{m}{k} x^k$	$m \in \mathbb{N}, R = +\infty$
$(1+x)^\alpha = 1 + \sum_{n=1}^{+\infty} \frac{\alpha(\alpha-1) \cdots (\alpha-n+1)}{n!} x^n$	$\alpha \in \mathbb{R} \setminus \mathbb{N}, R = 1.$

Dans le repère $(O, \vec{i}, \vec{j}, \vec{k})$, nous plongeons dans un champ magnétique harmonique uniforme, noté $\vec{B}$, une plaque de métal d'épaisseur $2a$, épaisseur faible par rapport à ses deux autres dimensions. Notons que $\vec{B}$ est parallèle à la plaque et que $\vec{B} = B_0 e^{i\omega t} \vec{k}$, où i est l'imaginaire pur tel que $i^2 = -1$ et ω représente la pulsation. À chaque instant à l'intérieur du métal, le champ magnétique, noté $\vec{B}_1$, est uniforme et est égal à $\vec{B}_1 = B_{int} e^{i\omega t} \vec{k}$. Les variations de $\vec{B}_1$ vont générer des courants de Foucault. Le circuit magnétique est parcouru par des boules de densité de courant qui seront responsables de la création d'un champ magnétique correctif $\vec{B}_2$ qui s'ajoutera au champ magnétique initial $\vec{B}_1$. Les variations de $\vec{B}_2$ vont aussi générer des courants de Foucault. Il y a création d'un nouveau champ magnétique $\vec{B}_3$ etc. Nous montrons que $\vec{B}_2 = i \frac{w\sigma\mu}{2} x^2 B_{int} e^{i\omega t} \vec{k}$, où μ représente la perméabilité absolue du matériau et σ la conductivité du matériau. Ainsi nous avons :

$$\vec{B}_2 = i \frac{w\sigma\mu}{2} x^2 \vec{B}_1, \quad \vec{B}_3 = -\frac{(w\sigma\mu)^2}{2 \times 3 \times 4} x^4 \vec{B}_1, \quad \vec{B}_4 = i \frac{(w\sigma\mu)^3}{2 \times 3 \times 4 \times 5 \times 6} x^6 \vec{B}_1.$$

1. Montrer que le champ magnétique total $\vec{B}_T$ dans la plaque est égal au produit de $\vec{B}_1$ par une série infinie. Nous posons : $\delta = \sqrt{\frac{2}{w\sigma\mu}}$ qui représente l'épaisseur de peau du conducteur.

Rappelons que le DSE(0) de ch est égal à $\text{ch}(x) = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \dots + \frac{x^{2n}}{(2n)!} + \dots$ avec $R = +\infty$.

2. Le champ magnétique est continu à la traversée de la plaque métallique. Calculer $\vec{B}_T$ en fonction de $\vec{B}_0$.

3. Calculer la norme B_T de $\vec{B}_T$.

1. Nous avons : $\frac{1}{\delta^2} = \frac{w\sigma\mu}{2}$, or

$$\vec{B}_T = \vec{B}_1 \left(1 + \frac{i w \sigma \mu x^2}{2} - \frac{(w \sigma \mu)^2 x^4}{2 \times 3 \times 4} + \frac{i (w \sigma \mu)^3 x^6}{2 \times 3 \times 4 \times 5 \times 6} + \dots \right).$$

Donc, comme $i = \left(\frac{1+i}{\sqrt{2}} \right)^2$, nous avons :

$$\vec{B}_T = \vec{B}_1 \left(1 + \frac{1}{2!} \left((1+i) \frac{x}{\delta} \right)^2 + \frac{1}{4!} \left((1+i) \frac{x}{\delta} \right)^4 + \dots \right).$$

En conclusion : $\vec{B}_T = \vec{B}_1 \text{ch} \left((1+i) \frac{x}{\delta} \right)$.

2. À la surface de la plaque, nous prendrons $x = a$, nous avons alors $B_0 e^{i\omega t} = B_{int} e^{i\omega t} \text{ch} \left((1+i) \frac{a}{\delta} \right)$ car le champ magnétique est continu à la traversée de la plaque métallique.

Donc, $B_{int} = \frac{B_0}{\text{ch} \left((1+i) \frac{a}{\delta} \right)}$ et donc :

$$\vec{B}_T = B_0 e^{i\omega t} \frac{\text{ch} \left((1+i) \frac{x}{\delta} \right)}{\text{ch} \left((1+i) \frac{a}{\delta} \right)} \vec{k}.$$

3. Nous avons $\|B_T\| = \sqrt{B_T B_T^*}$ où B_T, B_T^* sont $B_0 e^{i\omega t} \operatorname{ch}\left((1+i)\frac{x}{\delta}\right) / \operatorname{ch}\left((1+i)\frac{a}{\delta}\right)$ et sont conjugué. Or $\operatorname{ch}\left((1+i)\frac{x}{\delta}\right) = \frac{e^{\frac{(1+i)x}{\delta}} + e^{-\frac{(1+i)x}{\delta}}}{2}$ donc,

$$\operatorname{ch}\left((1+i)\frac{x}{\delta}\right) \times \operatorname{ch}\left((1-i)\frac{x}{\delta}\right) = \frac{e^{\frac{2x}{\delta}} + e^{-\frac{2x}{\delta}} + e^{\frac{2ix}{\delta}} + e^{-\frac{2ix}{\delta}}}{4} = \frac{\cos\left(\frac{2x}{\delta}\right) + \operatorname{ch}\left(\frac{2x}{\delta}\right)}{2} \in \mathbb{R}$$

$$\operatorname{ch}\left((1+i)\frac{a}{\delta}\right) \times \operatorname{ch}\left((1-i)\frac{a}{\delta}\right) = \frac{e^{\frac{2a}{\delta}} + e^{-\frac{2a}{\delta}} + e^{\frac{2ia}{\delta}} + e^{-\frac{2ia}{\delta}}}{4} = \frac{\cos\left(\frac{2a}{\delta}\right) + \operatorname{ch}\left(\frac{2a}{\delta}\right)}{2} \in \mathbb{R}$$

et donc :

$$B_T = B_0 \sqrt{\frac{\left|\cos\left(\frac{2x}{\delta}\right) + \operatorname{ch}\left(\frac{2x}{\delta}\right)\right|}{\left|\cos\left(\frac{2a}{\delta}\right) + \operatorname{ch}\left(\frac{2a}{\delta}\right)\right|}}.$$

Pour s'entraîner (solutions p. 529)

51.1 Exprimer $\frac{1}{1+x^2}$ comme la somme d'une série entière et donner l'intervalle de convergence.

51.2 Exprimer $\frac{1}{2+x}$ comme la somme d'une série entière et donner l'intervalle de convergence.

51.3 Exprimer $\cos(2x)$ comme la somme d'une série entière et donner l'intervalle de convergence.

51.4 Exprimer $\operatorname{sh}\left(\frac{x}{2}\right)$ comme la somme d'une série entière et donner l'intervalle de convergence.

51.5 Exprimer $\ln(1+x)$ comme la somme d'une série entière et donner l'intervalle de convergence.

51.6 Exprimer $\frac{x^3}{2+x}$ comme la somme d'une série entière et donner l'intervalle de convergence.

51.7 Exprimer $\ln(x^2 - 5x + 6)$ comme la somme d'une série entière et donner l'intervalle de convergence.

51.8 Exprimer $\ln(1+x-2x^2)$ comme la somme d'une série entière et donner l'intervalle de convergence.

51.9 Exprimer $(\arcsin(x))^2$ comme la somme d'une série entière et donner l'intervalle de convergence.

51.10 Exprimer $\sqrt{x + \sqrt{x^2 + 1}}$ comme la somme d'une série entière et donner l'intervalle de convergence.

51.11 Calculer l'intervalle de convergence et la somme de la série entière $\sum_{n=0}^{+\infty} a_n x^n$. La suite $(a_n)_{n \in \mathbb{N}}$ étant définie par ses premiers termes a_0, a_1, a_2 et la relation : $a_n = a_{n-1} + a_{n-2} - a_{n-3}$.

51.12 Calculer $\sum_{n=0}^{+\infty} \frac{x^{3n}}{(3n)!}$.

51.13 Calculer l'intervalle de convergence et la somme de la série entière $\sum_{n=1}^{+\infty} \frac{x^n}{n} \cos\left(\frac{2n\pi}{3}\right)$.

Définition 52.1

Soient f une fonction T -périodique, continue par morceaux et α un réel. Nous appelons **coefficients de Fourier réels** de f , les nombres réels définis, pour tout $n \in \mathbb{N}$, par :

$$a_n = \frac{2}{T} \int_{\alpha}^{\alpha+T} f(t) \cos\left(\frac{2\pi nt}{T}\right) dt, \quad b_n = \frac{2}{T} \int_{\alpha}^{\alpha+T} f(t) \sin\left(\frac{2\pi nt}{T}\right) dt.$$

Nous appelons **coefficients de Fourier complexes** de f , les nombres complexes définis, pour tout $n \in \mathbb{Z}$, par :

$$c_n = \frac{1}{T} \int_{\alpha}^{\alpha+T} e^{-\frac{i2\pi nt}{T}} f(t) dt.$$

Définition 52.2

Soient $T > 0$ et f une fonction T -périodique, continue par morceaux. Nous associons à f une série appelée **série de Fourier** qui, lorsqu'elle converge, définit une fonction S périodique de période T et qui est égale, sous forme réelle à :

$$\forall t \in \mathbb{R}, \quad S(t) = \frac{a_0}{2} + \sum_{n=1}^{+\infty} a_n \cos\left(\frac{2\pi nt}{T}\right) + b_n \sin\left(\frac{2\pi nt}{T}\right),$$

ou, sous forme complexe à :

$$\forall t \in \mathbb{R}, \quad S(t) = \sum_{n=-\infty}^{+\infty} c_n e^{\frac{i2\pi nt}{T}}.$$

Théorème 52.1

- Si la fonction f est paire, alors pour tout $n \in \mathbb{N}$, $b_n = 0$ ou encore $c_n = c_{-n}$.
- Si la fonction f est impaire, alors pour tout $n \in \mathbb{N}$, $a_n = 0$ ou encore $c_n = -c_{-n}$.
- Si en plus de l'une ou de l'autre des deux propriétés énoncées ci-dessus, nous avons :

$$\forall x \in \mathbb{R}, f\left(x + \frac{T}{2}\right) = f(x) \text{ alors } \forall k \in \mathbb{N}, a_{2k+1} = 0 \text{ ou } b_{2k+1} = 0, a_{2k} = \frac{8}{T} \int_{\alpha}^{\frac{T}{4}+\alpha} f(x) \cos\left(\frac{2\pi(2k)x}{T}\right) dx \text{ si } f(x) \text{ est paire.}$$

$$b_{2k} = \frac{8}{T} \int_{\alpha}^{\frac{T}{4}+\alpha} f(x) \sin\left(\frac{2\pi(2k)x}{T}\right) dx \text{ si } f(x) \text{ est impaire.}$$

- Si en plus de l'une ou de l'autre des deux propriétés énoncées ci-dessus, nous avons :

$$\forall x \in \mathbb{R}, f\left(x + \frac{T}{2}\right) = -f(x) \text{ alors } \forall k \in \mathbb{N}, a_{2k} = 0 \text{ ou } b_{2k} = 0,$$

$$\text{et } a_{2k+1} = \frac{8}{T} \int_{\alpha}^{\frac{T}{4}+\alpha} f(x) \cos\left(\frac{2\pi(2k+1)x}{T}\right) dx \text{ si } f(x) \text{ est paire.}$$

$$b_{2k+1} = \frac{8}{T} \int_{\alpha}^{\frac{T}{4}+\alpha} f(x) \sin\left(\frac{2\pi(2k+1)x}{T}\right) dx \text{ si } f(x) \text{ est impaire.}$$

Nous donnons ci-dessous des théorèmes sur la convergence de la série de Fourier d'une fonction.

Théorème 52.2 Théorème de Riemann-Lebesgue

Si f est une fonction d'un segment $J = [a; b]$ à valeurs complexes et intégrable sur J alors

$$\lim_{n \rightarrow +\infty} \int_a^b f(t) e^{-int} dt = 0.$$

Corollaire 52.1

Les coefficients de Fourier d'une fonction périodique et intégrable sur une période tendent vers 0 lorsque n tend vers l'infini.

Théorème 52.3 Théorème de Dirichlet

Soit f une fonction, définie sur $\mathbb{R}$ et à valeurs complexes, T -périodique de classe $\mathcal{C}^1$ par morceaux. Notons $f(x^-)$ et $f(x^+)$ les limites à gauche et à droite de f au point x réel. La série de Fourier associée à $f(x)$ est simplement convergente sur $\mathbb{R}$ et l'égalité suivante est vérifiée en tout point x réel :

$$\frac{f(x^+) + f(x^-)}{2} = \frac{a_0}{2} + \sum_{n=1}^{+\infty} a_n \cos\left(\frac{2\pi nx}{T}\right) + b_n \sin\left(\frac{2\pi nx}{T}\right)$$

ou encore

$$\frac{f(x^+) + f(x^-)}{2} = c_0 + \sum_{n=1}^{+\infty} c_n e^{\left(\frac{2i\pi nx}{T}\right)} + c_{-n} e^{\left(\frac{2i\pi nx}{T}\right)}.$$

Ainsi, en tout point où f est continue, nous avons :

$$f(x) = \sum_{n=0}^{+\infty} a_n \cos\left(\frac{2\pi nx}{T}\right) + b_n \sin\left(\frac{2\pi nx}{T}\right)$$

ou encore

$$f(x) = c_0 + \sum_{n=1}^{+\infty} c_n e^{\left(\frac{2i\pi nx}{T}\right)} + c_{-n} e^{\left(\frac{2i\pi nx}{T}\right)}.$$

Théorème 52.4 Coefficients de Fourier de la dérivée

Soit f une fonction, définie sur $\mathbb{R}$ et à valeurs complexes, T -périodique de classe $\mathcal{C}^1$ par morceaux et continue dont les coefficients de Fourier sont notés c_n . Les coefficients de Fourier c'_n de la dérivée f' de f sont donnés par $c'_n = inc_n$.

Théorème 52.5 Convergences absolue et uniforme

Soit f une fonction, définie sur $\mathbb{R}$ et à valeurs complexes, T -périodique de classe $\mathcal{C}^1$ par morceaux et continue. La série de Fourier de f converge absolument et uniformément vers f .

Théorème 52.6 Formule de Parseval

Soient f une fonction T -périodique et intégrable et α un réel. Ses coefficients de Fourier complexes et réels vérifient les relations suivantes :

$$\frac{1}{T} \int_{\alpha}^{\alpha+T} |f(x)|^2 dx = \sum_{n=-\infty}^{+\infty} |c_n|^2 \text{ avec } |c_n|^2 = |c_{-n}|^2 = \frac{a_n^2 + b_n^2}{4}.$$

Nous avons aussi :

$$\frac{1}{T} \int_{\alpha}^{T+\alpha} |f(x)|^2 dx = a_0^2 + \sum_{n=1}^{+\infty} \frac{a_n^2 + b_n^2}{2} \text{ avec } a_0 = \frac{1}{T} \int_{\alpha}^{T+\alpha} f(x) dx.$$

Application

1. Tension canal de transmission

Soit $u(t)$ une tension de forme carré, d'amplitude E , T -périodique représentée à la Figure 52.

a. Calculer les coefficients de la série de Fourier de $u(t)$.

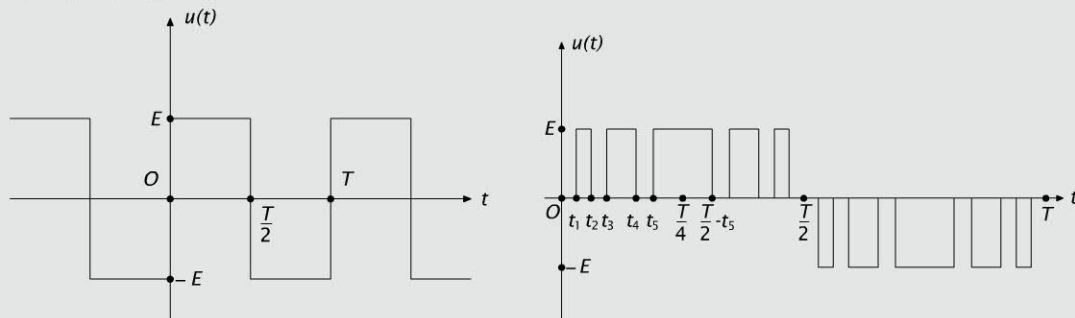
b. Nous ne pouvons transmettre que 90% de la puissance de ce signal à cause de la bande passante du canal de transmission. Combien d'harmoniques pouvons-nous conserver pour que la transmission soit correcte ?

2. Moteur électrique

Pour le fonctionnement correct d'un système comme un moteur électrique avec son électronique de puissance, il est nécessaire d'éliminer des harmoniques de faible rang au-dessus du fondamental de fréquence f qui occasionnent des pertes énergétiques et un dysfonctionnement du moteur. Par exemple, pour éliminer les harmoniques de rang $3f, 5f, 7f, 9f$ et $11f$ dans une tension rectangulaire appliquée à l'électronique de puissance d'un moteur électrique, nous allons introduire des fentes dans ce signal rectangulaire. Nous prendrons $T = 10^{-2}$ s.

a. Calculer la série de Fourier de cette tension $u(t)$, représentée sur la Figure 52.

b. En déduire les valeurs de t_1, t_2, t_3, t_4 et t_5 pour éliminer les harmoniques de rang : $3f, 5f, 7f, 9f, 11f$.



Courbes des Applications 1 (à gauche) et 2 (à droite).

1.a. Remarquons que $u(-t) = -u(t)$. Le signal est donc impair et ainsi pour tout $n \geq 0$, nous avons $a_n = 0$. De plus, notons que $u\left(t + \frac{T}{2}\right) = -u(t)$, ce qui entraîne $b_{2k} = 0$ et

$$b_{2k+1} = \frac{8}{T} \int_0^{\frac{T}{4}} E \sin\left(\frac{2\pi(2k+1)t}{T}\right) dt = \frac{8ET}{T(2\pi(2k+1))} \left(-\cos\left(\frac{\pi(2k+1)}{2}\right) + 1\right) = \frac{4E}{\pi(2k+1)}.$$



La valeur moyenne du signal est nulle ce qui corrobore bien que $a_0 = \frac{1}{T} \int_0^T u(t) dt = 0$.

Donc, en posant $f = \frac{1}{T}$, nous avons $u(t) = \frac{4E}{\pi} \sum_{k=0}^{+\infty} \frac{1}{2k+1} \sin(2\pi(2k+1)ft)$.



La composante de fréquence f s'appelle le **fondamental du signal**, les autres composantes de fréquences $3f, 5f, 7f, \dots$ s'appellent les **harmoniques du signal**.

b. Rappelons que la valeur efficace du signal est égale à $U_{eff} = \sqrt{\frac{1}{T} \int_0^T u^2(t) dt}$, et la puissance du signal est égale à $P = U_{eff}^2 = \frac{2}{T} \int_0^{\frac{T}{2}} E^2 dt = E^2$. Donc nous voulons transmettre $P_T = 0,9E^2$. Le théorème de Parseval donne $U_{eff}^2 = \sum_{k=0}^{+\infty} \frac{b_{2k+1}^2}{2} = E^2 \times 0,81 \times \left(1 + \frac{1}{9} + \frac{1}{25} + \dots\right)$. Nous constatons que $0,81 \times \left(1 + \frac{1}{9}\right) = 0,9$. Donc pour avoir 90 % de la puissance, il suffit de transmettre le fondamental de fréquence f et l'harmonique de fréquence $3f$ donc deux harmoniques nécessaires.

2.a. Nous étudions les symétries : $u(-t) = -u(t)$ donc $a_n = 0$, $u(t + \frac{T}{2}) = -u(t)$ donc $b_{2k} = 0$. Donc, nous n'avons que des harmoniques de rang impair : b_{2k+1} .

$$b_{2k+1} = \frac{8}{T} \int_0^{\frac{T}{4}} u(t) \sin\left(\frac{2\pi(2k+1)t}{T}\right) dt = \frac{8}{T} \left(\int_{t_1}^{t_2} E \sin\left(\frac{2\pi(2k+1)t}{T}\right) dt + \int_{t_3}^{t_4} E \sin\left(\frac{2\pi(2k+1)t}{T}\right) dt + \int_{t_5}^{\frac{T}{4}} E \sin\left(\frac{2\pi(2k+1)t}{T}\right) dt \right).$$

$$b_{2k+1} = \frac{4E}{\pi(2k+1)} \left(-\cos\left(\frac{2\pi(2k+1)t_2}{T}\right) + \cos\left(\frac{2\pi(2k+1)t_1}{T}\right) - \cos\left(\frac{2\pi(2k+1)t_4}{T}\right) + \cos\left(\frac{2\pi(2k+1)t_3}{T}\right) + \cos\left(\frac{2\pi(2k+1)t_5}{T}\right) \right).$$

b. Nous voulons éliminer cinq harmoniques donc il faut cinq équations, donc pour $k = 1, 2, 3, 4, 5$: $-\cos\left(\frac{2\pi(2k+1)t_2}{T}\right) + \cos\left(\frac{2\pi(2k+1)t_1}{T}\right) - \cos\left(\frac{2\pi(2k+1)t_4}{T}\right) + \cos\left(\frac{2\pi(2k+1)t_3}{T}\right) + \cos\left(\frac{2\pi(2k+1)t_5}{T}\right) = 0$.

Avec Maple, nous obtenons : $t_1 = 2,89 \cdot 10^{-2} \text{ s}$, $t_2 = 4,24 \cdot 10^{-2} \text{ s}$, $t_3 = 5,87 \cdot 10^{-2} \text{ s}$, $t_4 = 8,42 \cdot 10^{-2} \text{ s}$, $t_5 = 9,03 \cdot 10^{-2} \text{ s}$.

Pour s'entraîner

52.1

www

1. Soient $f(x) = \cos(ax)$ pour $x \in [-\pi; \pi]$ et $a \in \mathbb{R} \setminus \mathbb{Z}$. Développer f en série de Fourier. En déduire la valeur des sommes : $\sum_{n=1}^{+\infty} \frac{(-1)^n}{a^2 - n^2}$, $\sum_{n=1}^{+\infty} \frac{1}{a^2 - n^2}$ et $\sum_{n=1}^{+\infty} \frac{1}{(a^2 - n^2)^2}$.

2. Même question avec $g(x) = \text{ch}(ax)$ pour $x \in [-\pi; \pi]$ et $a \in \mathbb{R}^*$. En déduire la valeur des sommes : $\sum_{n=1}^{+\infty} \frac{(-1)^n}{n^2 + a^2}$, $\sum_{n=1}^{+\infty} \frac{1}{n^2 + a^2}$ et $\sum_{n=1}^{+\infty} \frac{1}{(n^2 + a^2)^2}$.

52.2

www

Soit $S(x) = \sum_{n \geq 1} u_n(x)$ avec $u_n(x) = (-1)^n \frac{\cos(nx)}{n^2(n^2 + a^2)}$ et a réel fixé.

1. Montrer que S est de classe $\mathcal{C}^2$ sur $\mathbb{R}$ et vérifie sur $[-\pi; \pi]$ l'équation différentielle :

$$S''(x) - a^2 S(x) = \frac{\pi^2 - 3x^2}{12}.$$

2. En déduire $S(x)$.

Intégration sur un intervalle quelconque

1. Intégrale convergente, intégrale divergente

Définition 53.1

Soit f une fonction définie sur un intervalle I de $\mathbb{R}$. La fonction f est dite **localement intégrable sur l'intervalle I** si elle est continue par morceaux sur I et intégrable sur tout segment inclus dans I .

Définition 53.2

Soit f une fonction localement intégrable sur l'intervalle $]a; b]$ avec $a < b$, où $a \in \mathbb{R} \cup \{-\infty\}$. Si $\lim_{x \rightarrow a^+} \int_x^b f(t)dt$ existe, nous disons que $\int_a^b f(t)dt$, **parfois aussi notée $\int_I f(t)dt$, converge ou est convergente**. Dans le cas contraire, nous disons que **l'intégrale diverge ou est divergente**.

Lorsque l'intégrale $\int_I f(t)dt$ est convergente, sa valeur est égale à $\lim_{x \rightarrow a^+} \int_x^b f(t)dt$.



- Pour définir une intégrale généralisée sur $I = [a; b[$, avec $b \in \mathbb{R} \cup \{+\infty\}$, ou $I =]a; b[$, avec $a \in \mathbb{R} \cup \{-\infty\}$ et $b \in \mathbb{R} \cup \{+\infty\}$, nous procédons de manière similaire.
- Étudier la nature d'une intégrale généralisée, c'est déterminer si elle est convergente ou divergente. Nous disons que deux intégrales sont de même nature si elles sont toutes deux convergentes ou toutes deux divergentes.

2. Règles de convergence

Théorème 53.1

Soient f et g deux fonctions continues par morceaux sur un intervalle I , si $\int_I f(t)dt$ et $\int_I g(t)dt$ convergent alors $\int_I (f(t) + g(t))dt$ converge.

Si $\int_I f(t)dt$ converge et $\int_I g(t)dt$ diverge alors $\int_I (f(t) + g(t))dt$ diverge.

Si $\int_I f(t)dt$ et $\int_I g(t)dt$ divergent alors nous ne pouvons rien dire de la nature de $\int_I (f(t) + g(t))dt$.

Théorème 53.2

Soit f une fonction localement intégrable et positive sur I . Si $\int_I f(t)dt$ converge alors $\int_I f(t)dt \geq 0$.

Lorsque $I = [a; b[$, avec $b \in \mathbb{R} \cup \{+\infty\}$, $\int_I f(t)dt$ converge si et seulement si la fonction φ définie par : $\varphi(x) = \int_a^x f(t)dt$ est majorée sur $[a; b[$, et $\int_a^b f(t)dt$ diverge si et seulement si $\lim_{x \rightarrow b} \varphi(x) \rightarrow +\infty$.

Proposition 53.1 Comparaison de fonctions positives

Soient f et g deux fonctions localement intégrables et telles que $0 \leq f \leq g$ sur I .

Si $\int_I g(t)dt$ converge, alors $\int_I f(t)dt$ converge aussi et $\int_I f(t)dt \leq \int_I g(t)dt$.

Si $\int_I f(t)dt$ diverge, alors $\int_I g(t)dt$ diverge aussi.

Proposition 53.2 Comparaison de fonctions positives

Soient f et g deux fonctions localement intégrables et positives sur $I = [a; b[$, avec $b \in \mathbb{R} \cup \{+\infty\}$. Si $f(x) \sim_b g(x)$, alors $\int_a^b f(t)dt$ et $\int_a^b g(t)dt$ sont de même nature, c'est-à-dire qu'elles sont toutes les deux convergentes ou toutes les deux divergentes. La notation $f(x) \sim_b g(x)$ signifie « f et g sont équivalentes en b », i.e. $\lim_{x \rightarrow b} (f(x) - g(x))/g(x) = 0$.



Il est important que f et g soient de même signe au voisinage du problème étudié, sinon les fonctions peuvent être équivalentes et leurs intégrales de natures différentes.

Théorème 53.3 Intégrales de Riemann

1. Soit $a > 0$. $\int_a^{+\infty} \frac{dt}{t^\alpha}$ converge si et seulement si $\alpha > 1$.
2. Soit $(a, b) \in \mathbb{R}^2$ avec $a < b$, $\int_a^b \frac{dt}{(b-t)^\alpha}$ converge si et seulement si $\alpha < 1$.

Théorème 53.4 Critère de Riemann

1. Soit f une fonction localement intégrable et positive sur $]0; a]$ avec $a > 0$. Nous montrons que $\int_0^a f(t)dt$ converge en trouvant deux réels $c \leq a$ et $\alpha < 1$ tels que pour tout $x \in]0; c]$, $f(x) \leq \frac{1}{x^\alpha}$ ou en trouvant $\alpha < 1$ tel que $\lim_{x \rightarrow 0} x^\alpha f(x) = l \in \mathbb{R}$.
2. Soit f une fonction localement intégrable et positive sur $[a; +\infty[$ avec $a > 0$. Nous montrons que $\int_a^{+\infty} f(t)dt$ converge en trouvant deux réels $c \geq a$ et $\alpha > 1$ tels que pour tout $x \in [c; +\infty[$, $f(x) \leq \frac{1}{x^\alpha}$ ou en trouvant $\alpha > 1$ tel que $\lim_{x \rightarrow +\infty} x^\alpha f(x) = l \in \mathbb{R}$.

Théorème 53.5 Intégrales de Bertrand

$\forall a > 1$, $\forall (\alpha, \beta) \in \mathbb{R}^2$, $\int_a^{+\infty} \frac{1}{t^\alpha (\ln t)^\beta} dt$ converge si et seulement si $\alpha > 1$ ou $\alpha = 1$ et $\beta > 1$.

3. Intégrales absolument convergentes**Définition 53.3**

Soit f une fonction localement intégrable sur I . Nous disons que l'intégrale de f est **absolument convergente** sur cet intervalle si $\int_I |f(t)|dt$ converge.

Théorème 53.6

Soit f une fonction localement intégrable sur I , si l'intégrale de f est absolument convergente sur I , alors $\int_I f(t)dt$ est convergente et nous avons : $|\int_I f(t)dt| \leq \int_I |f(t)|dt$.



1. Champ électrostatique

L'espace $\mathbb{R}^3$ est rapporté au repère orthonormé direct $(O, \vec{u}(\theta), \vec{v}(\theta), \vec{k})$ des coordonnées polaires. Soit le champ électrostatique $\vec{E}$ créé en un point M quelconque de l'espace par une distribution de charges de densité linéique λ uniforme et répartie le long de l'axe dirigé par $\vec{k}$. La composante du champ $\vec{E}$:

1. suivant le vecteur $\vec{u}(\theta)$ est $E_r = \int_{-\infty}^{+\infty} \frac{\lambda r dz}{4\pi\epsilon_0 (\sqrt{r^2 + z^2})^3}$;
2. suivant le vecteur $\vec{k}$ est $E_z = \int_{-\infty}^{+\infty} \frac{-\lambda z dz}{4\pi\epsilon_0 (\sqrt{r^2 + z^2})^3}$.

Calculer E_z et E_r .

2. Fonction Gamma

La fonction Gamma, notée Γ , est définie pour tout $x > 0$ par :

$$\Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t} dt.$$

Montrer que Γ est définie sur $\mathbb{R}_+^*$ et que, $\forall x \in \mathbb{R}_+^*$, $\Gamma(x+1) = x\Gamma(x)$ et, $\forall n \in \mathbb{N}^*$, $\Gamma(n) = (n-1)!$.

3. Études de convergence

Étudier la convergence de :

$$\text{a. } I_1 = \int_1^{+\infty} \frac{\sin(x)dx}{x^2}; \quad \text{b. } I_2 = \int_0^{\pi} \frac{\sin(x)dx}{x^{\frac{3}{2}}}; \quad \text{c. } I_3 = \int_0^1 \frac{x^5 dx}{\sqrt{1-x^4}}.$$

Nous posons $\tan(\alpha) = \frac{z}{r}$, avec $\alpha \in]-\frac{\pi}{2}; \frac{\pi}{2}[$ et donc $d(\tan(\alpha)) = (1 + \tan^2(\alpha)) d\alpha = \frac{d\alpha}{\cos^2(\alpha)} = \frac{dz}{r}$.

Lorsque $z \rightarrow -\infty$, $\alpha \rightarrow -\frac{\pi}{2}$ et lorsque $z \rightarrow +\infty$, $\alpha \rightarrow \frac{\pi}{2}$.

Nous avons : $z = r \tan(\alpha)$, $dz = \frac{r d\alpha}{\cos^2(\alpha)}$, $\sqrt{r^2 + z^2} = \frac{r}{\cos(\alpha)}$ car $\cos(\alpha) > 0$. Par conséquent :

$$E_z = \frac{-\lambda}{4\pi\epsilon_0 r} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sin(\alpha) d\alpha = \frac{\lambda}{4\pi\epsilon_0 r} [\cos(\alpha)]_{-\frac{\pi}{2}}^{\frac{\pi}{2}} = 0$$

et

$$E_r = \frac{\lambda}{4\pi\epsilon_0 r} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos(\alpha) d\alpha = \frac{\lambda}{4\pi\epsilon_0 r} [\sin(\alpha)]_{-\frac{\pi}{2}}^{\frac{\pi}{2}} = \frac{\lambda}{2\pi\epsilon_0 r}.$$

2. La fonction $t \mapsto t^{x-1} e^{-t}$ est continue, positive sur $]0; +\infty[$. Si $t \rightarrow 0$, $e^{-t} \rightarrow 1$ et $t^{x-1} e^{-t}$ équivalent autour de 0 à $\frac{1}{t^{1-x}}$.

Les fonctions étant positives, les intégrales $\int_0^1 t^{x-1} e^{-t} dt$ et $\int_0^1 \frac{dt}{t^{1-x}}$ sont de même nature donc elles convergent si et seulement si $x > 0$.

La fonction $t \mapsto t^{x+1} e^{-t}$ est continue sur $[1; +\infty[$, elle tend vers 0 quand $t \rightarrow +\infty$ donc quand $t \rightarrow +\infty$, $t^{x-1} e^{-t} = o\left(\frac{1}{t^2}\right)$. Le critère de Riemann permet donc d'affirmer que $\Gamma(x)$ est définie $\forall x > 0$.

3. a. Soit $f(x) = \frac{\sin(x)}{x^2}$. La fonction f est définie sur $[1; +\infty[$, f est continue sur $[1; +\infty[$ et $\forall x \in [1; +\infty[$, nous avons : $|f(x)| \leq \frac{1}{x^2}$. Le critère de Riemann (Théorème 53.4) permet alors d'affirmer que I_1 est absolument convergente donc convergente.

b. Soit $f(x)$ la fonction de $]0; \pi]$ dans $\mathbb{R}$ définie par : $f(x) = \frac{\sin(x)}{x^{\frac{3}{2}}}$, f est positive et continue sur $]0; \pi]$ et $\lim_{x \rightarrow 0^+} x^{\frac{1}{2}}|f(x)| = \lim_{x \rightarrow 0^+} \frac{|\sin(x)|}{x} = \lim_{x \rightarrow 0^+} \frac{\sin(x)}{x} = \sin'(0) = \cos(0) = 1 \in \mathbb{R}$. Le critère de Riemann (Théorème 53.4) permet alors d'affirmer que I_2 est absolument convergente donc convergente.

c. Soit $f(x) = \frac{x^5}{\sqrt{1-x^4}}$. f est continue et positive sur $[0; 1[$ et $\forall x \in [0; 1[$, nous avons : $0 \leq \frac{x^5}{\sqrt{1-x^4}} = \frac{x^5}{\sqrt{1-x}\sqrt{1+x}\sqrt{1+x^2}} \leq \frac{1}{\sqrt{1-x}}$. Or, $\int_0^1 \frac{dx}{\sqrt{1-x}}$ converge donc la Proposition 53.1 permet d'affirmer que I_3 converge.

Pour s'entraîner (solutions p. 530)

53.1 Étudier la nature des intégrales :

1. $\int_0^1 \frac{dx}{e^x - \cos x}$
2. $\int_1^{+\infty} \frac{\arctan(x-1)}{(x^2-1)^{\frac{4}{3}}} dx$
3. $\int_{\frac{1}{2}}^1 \frac{dx}{x(1-x^3)^{\frac{2}{3}}}$
4. $\int_0^{+\infty} \left(\frac{\sin x}{x}\right)^{\frac{6}{5}} dx$

53.2  **Quelques intégrales généralisées à connaître :**

1. $\int_a^{+\infty} \frac{dx}{x^2} \quad (a > 0)$
2. $\int_0^1 \ln(x) dx$
3. $\int_a^{+\infty} \exp(-\alpha x) dx \quad (a \in \mathbb{R}, \alpha > 0)$
4. $\int_1^{+\infty} \frac{\ln x}{x^2} dx$
5. $\int_0^{+\infty} \frac{\ln x}{(1+x)^2} dx$
6. $\int_0^{+\infty} x^n \exp(-x) dx \quad (n \in \mathbb{N})$
7. $\int_a^{+\infty} \frac{dx}{x(x+r)} \quad (a > 0, r > 0)$
8. $\int_0^{+\infty} \frac{\arctan x}{1+x^2} dx$

53.3 $f \sim g$ et pourtant leurs intégrales ne sont pas de même nature

Soit $a > 0$. Nous définissons sur $[a; +\infty[$, les fonctions f et g par :

$$f(x) = \frac{\sin x}{\sqrt{x}} + \frac{\sin^2 x}{x} \quad \text{et} \quad g(x) = \frac{\sin x}{\sqrt{x}}.$$

1. Montrer que $f \sim g$ au voisinage de $+\infty$.

2. Montrer que les intégrales $\int_a^{+\infty} f(x) dx$ et $\int_a^{+\infty} g(x) dx$ n'ont pas la même convergence.

53.4  **Un classique**

$$\text{Soit } I = \int_0^{+\infty} \frac{\exp(-x) - \exp(-2x)}{x} dx.$$

1. Montrer que I est convergente.

2. Pour $\varepsilon > 0$, établir, en posant $t = 2x$, la relation

$$\int_{\varepsilon}^{+\infty} \frac{\exp(-x) - \exp(-2x)}{x} dx = \int_{\varepsilon}^{2\varepsilon} \frac{\exp(-x)}{x} dx.$$

3. En déduire la valeur de I .

53.5 **Une astuce**

$$\text{Soit } J = \int_0^{\pi/2} \ln(\sin x) dx.$$

1. Montrer que J est convergente et que nous avons $J = \int_0^{\pi/2} \ln(\cos x) dx$.

2. Montrer que $2J = \int_0^{\pi/2} \ln\left(\frac{\sin 2x}{2}\right) dx$.

3. En déduire la valeur de J .

Convergences monotone et dominée

Intégrales dépendant d'un paramètre

1. Convergences monotone et dominée

Définition 54.1

Une suite de fonctions intégrables sur un intervalle I **converge en moyenne** vers une fonction f intégrable sur I si $\int_I |f_n - f|$ tend vers 0 quand n tend vers $+\infty$.

Proposition 54.1

La convergence en moyenne d'une suite (f_n) de fonctions intégrables entraîne la convergence de la suite des intégrales :

$$\int_I f = \int_I \lim_{n \rightarrow +\infty} f_n = \lim_{n \rightarrow +\infty} \int_I f_n.$$

Théorème 54.1 Convergence bornée

Soit (f_n) une suite de fonctions à valeurs réelles ou complexes continues par morceaux sur le segment $J = [a; b]$, telle que (f_n) converge simplement sur J vers une fonction f continue par morceaux sur J , et qu'il existe une constante $M > 0$ vérifiant $|f_n(x)| \leq M$ pour tout entier n et tout $x \in J$, alors la suite (f_n) converge en moyenne vers f sur J .

Théorème 54.2 Convergence monotone

Soit (f_n) une suite croissante de fonctions à valeurs réelles continues par morceaux sur le segment J convergeant simplement sur J vers une fonction f continue par morceaux sur J , alors nous avons

$$\int_J f = \sup_{n \in \mathbb{N}} \int_J f_n = \lim_{n \rightarrow +\infty} \int_J f_n.$$

Théorème 54.3 Convergence dominée

Soit (f_n) une suite de fonctions à valeurs réelles ou complexes continues par morceaux sur l'intervalle I et ϕ une fonction continue par morceaux, positive et intégrable sur I . Si (f_n) converge simplement sur I vers une fonction f continue par morceaux sur I et si, pour tout entier n , nous avons la domination $|f_n| \leq \phi$, alors f est intégrable sur I et la suite (f_n) converge en moyenne vers f . Par conséquent, la formule d'interversion a lieu :

$$\int_I f = \lim_{n \rightarrow +\infty} \int_I f_n = \lim_{n \rightarrow +\infty} \int_I f_n.$$

Théorème 54.4 Convergence monotone

Soit (f_n) une suite croissante de fonctions à valeurs réelles continues par morceaux et intégrables sur l'intervalle I convergeant simplement sur I vers une fonction f continue par morceaux sur I . Alors f est intégrable sur I si et seulement si la suite $(\int_I f_n)$ est

majorée. Dans ces conditions, la suite (f_n) converge en moyenne vers f et nous avons

$$\int_I f = \sup_{n \in \mathbb{N}} \int_I f_n = \lim_{n \rightarrow +\infty} \int_I f_n.$$

Théorème 54.5 Intégration des séries de fonctions positives

Soit (u_n) une suite de fonctions continues par morceaux à valeurs réelles positives et intégrables sur l'intervalle I , telles que la série $\sum_n u_n$ converge simplement sur I , ayant pour somme une fonction u continue par morceaux sur I . Alors u est intégrable si et seulement si nous avons $\sum_{n=0}^{+\infty} \int_I u_n$ converge ; dans ce cas, nous avons

$$\int_I \sum_{n=0}^{+\infty} u_n = \sum_{n=0}^{+\infty} \int_I u_n.$$

Théorème 54.6 Intégration des séries de fonctions

Soit (u_n) une suite de fonctions continues par morceaux à valeurs réelles ou complexes et intégrables sur l'intervalle I telles que la série $\sum_n u_n$ converge simplement sur I vers une fonction S continue par morceaux sur I . Alors si la série $\sum_n \int_I |u_n|$ converge, S est intégrable sur I et

$$\int_I \left| \sum_{n=0}^{+\infty} u_n \right| = \sum_{n=0}^{+\infty} \int_I |u_n|, \quad \int_I \sum_{n=0}^{+\infty} u_n = \sum_{n=0}^{+\infty} \int_I u_n.$$

2. Intégrales dépendant d'un paramètre

Théorème 54.7 Continuité sous le signe $\int$

Soit f une fonction à valeurs réelles ou complexes définie sur $A \times I$, où A est une partie de $\mathbb{R}^m$ et I intervalle de $\mathbb{R}$, telle que

- (i) pour tout élément $x \in A$, la fonction $t \mapsto f(x, t)$ est continue par morceaux sur I ,
- (ii) pour tout élément $t \in I$, la fonction $x \mapsto f(x, t)$ est continue sur A ,
- (iii) il existe une fonction ϕ continue par morceaux, positive, intégrable sur I telle que $|f(x, t)| \leq \phi(t)$ pour tout $(x, t) \in A \times I$.

Alors la fonction g définie sur A par $g(x) = \int_I f(x, t) dt$ est continue.

Théorème 54.8 Dérivation sous le signe $\int$ de Leibniz

Soient A et I deux intervalles ouverts de $\mathbb{R}$ et f une fonction de $A \times I$ à valeurs réelles ou complexes telle que

- (i) pour tout élément $x \in A$, la fonction $t \mapsto f(x, t)$ est continue par morceaux et intégrable sur I ,
- (ii) pour tout élément $t \in I$, la fonction $x \mapsto f(x, t)$ est dérivable sur A ,
- (iii) pour tout élément $x \in A$, la fonction $t \mapsto \frac{\partial f}{\partial x}(x, t)$ est continue par morceaux et intégrable sur I ,
- (iv) il existe une fonction ϕ continue par morceaux, positive, intégrable sur I telle que $\left| \frac{\partial f}{\partial x}(x, t) \right| \leq \phi(t)$ pour tout $(x, t) \in A \times I$.

Alors la fonction g définie sur A par $g(x) = \int_I f(x, t) dt$ est dérivable sur A et

$$g'(x) = \int_I \frac{\partial f}{\partial x}(x, t) dt.$$

Si, de plus, $x \mapsto \frac{\partial f}{\partial x}(x, t)$ est continue sur A pour tout t fixé, alors g est de classe $\mathcal{C}^1$ sur A .

Théorème 54.9 Formule de Fubini d'intégration sous le signe $\int$ de Fichtenholz

Soit f une application de $[a; b] \times [c; d]$ dans $\mathbb{R}$ ou $\mathbb{C}$ telle que les applications partielles $x \mapsto f(x, y_0)$ et $y \mapsto f(x_0, y)$ soient continues par morceaux pour tout $x_0 \in [a; b]$ et $y_0 \in [c; d]$, et que les fonctions $y \mapsto \int_a^b f(x, y) dx$ et $x \mapsto \int_c^d f(x, y) dy$ soient continues par morceaux. Alors nous avons la formule :

$$\int_c^d \left(\int_a^b f(x, y) dx \right) dy = \int_a^b \left(\int_c^d f(x, y) dy \right) dx.$$

Théorème 54.10

Soit f une application de $[a; b] \times \mathbb{R}_+$ dans $\mathbb{R}$ ou $\mathbb{C}$ telle que

- les applications partielles $x \mapsto f(x, y_0)$ et $y \mapsto f(x_0, y)$ soient continues par morceaux pour tout $x_0 \in [a; b]$ et $y_0 \in \mathbb{R}_+$;
- les fonctions $y \mapsto \int_a^b f(x, y) dx$ et $x \mapsto \int_0^{+\infty} f(x, y) dy$ soient continues par morceaux ;
- il existe une fonction ϕ intégrable sur $\mathbb{R}^+$ telle que pour tout $x \in [a; b]$ et $y \geq 0$ nous ayons $|f(x, y)| \leq \phi(y)$.

Alors nous avons la formule :

$$\int_0^{+\infty} \left(\int_a^b f(x, y) dx \right) dy = \int_a^b \left(\int_0^{+\infty} f(x, y) dy \right) dx.$$

Application

Le rayonnement de Planck

La loi du rayonnement de Planck est donnée par la densité spectrale d'énergie :

$$D(\lambda, T) = \frac{2\pi c^2 h}{\lambda^5} \frac{1}{e^{\frac{hc}{\lambda kT}} - 1}.$$

Déterminer l'énergie par unité de volume, $U(T)$, qui se calcule à l'aide de la formule suivante :

$$U(T) = \int_0^{+\infty} D(\lambda, T) d\lambda = \int_0^{+\infty} \frac{2\pi c^2 h}{\lambda^5} \frac{d\lambda}{e^{\frac{hc}{\lambda kT}} - 1}.$$

Vous pourrez utiliser la relation suivante : $\sum_{n=1}^{+\infty} \frac{1}{n^4} = \frac{\pi^4}{90}$.



Effectuons le changement de variable $x = \frac{hc}{\lambda kT}$ avec $dx = \frac{-hc}{kT} \frac{d\lambda}{\lambda^2}$ qui est valide car la fonction $\lambda \mapsto \frac{hc}{\lambda kT}$ est bijective de $]0; +\infty[$ sur $]0; +\infty[$, et de classe $\mathcal{C}^1$ sur $]0; +\infty[$:

$$U(T) = 2\pi c^2 h \left(\frac{kT}{hc} \right)^4 \int_0^{+\infty} \frac{x^3 dx}{e^x - 1}.$$

Pour calculer $I = \int_0^{+\infty} \frac{x^3 dx}{e^x - 1}$, remarquons que, comme $x > 0$, la série $S(x) = \sum_{n=1}^{+\infty} e^{-nx}$ est une série géométrique convergente de premier terme e^{-x} et de somme $S(x) = \frac{1}{e^x - 1} = \frac{e^{-x}}{1 - e^{-x}}$.

En introduisant cette expression de S sous la forme d'une série de fonctions positives, nous obtenons : $I = \int_0^{+\infty} \sum_{n=1}^{+\infty} x^3 e^{-nx} dx$.



Il s'agit maintenant de justifier l'intervention entre la série $\sum_{n=1}^{+\infty} x^3 e^{-nx}$ et l'intégrale sur $[0; +\infty[$, ce qui nous permettra d'achever le calcul. Afin de pouvoir appliquer le Théorème 54.5, nous devons montrer que la série $T = \sum_{n=1}^{+\infty} \int_0^{+\infty} x^3 e^{-nx} dx$ converge.

Nous posons $nx = t$, qui est un changement de variable licite car affine avec $ndx = dt$, donc $T = \sum_{n=1}^{+\infty} \frac{1}{n^4} \int_0^{+\infty} t^3 e^{-t} dt = \alpha \sum_{n=1}^{+\infty} \frac{1}{n^4}$, avec $\alpha = \int_0^{+\infty} t^3 e^{-t} dt \in \mathbb{R}$. La série $\sum_{n=1}^{+\infty} \frac{1}{n^4}$ est un exemple de Riemann qui converge puisque $4 > 1$. Par conséquent :

$$\int_0^{+\infty} \frac{x^3 dx}{e^x - 1} = \int_0^{+\infty} \sum_{n=1}^{+\infty} x^3 e^{-nx} dx = \sum_{n=1}^{+\infty} \int_0^{+\infty} x^3 e^{-nx} dx = \int_0^{+\infty} t^3 e^{-t} dt \sum_{n=1}^{+\infty} \frac{1}{n^4}.$$

Pour calculer la valeur de $\int_0^{+\infty} t^3 e^{-t} dt$, utilisons plusieurs intégrations par parties successives ou la fonction Γ , en remarquant que $\int_0^{+\infty} t^3 e^{-t} dt = \Gamma(4) = 3! = 6$.

En utilisant le résultat fourni par l'énoncé, nous avons alors : $I = \frac{\pi^4}{90} \times 6 = \frac{\pi^4}{15}$. Nous en déduisons :

$$U(T) = \frac{2\pi^5 k^4 T^4}{15h^3 c^2}.$$

Fiche 53

Pour s'entraîner (solutions p. 530)

54.1 Étudier la suite (u_n) de terme général

$$u_n = \int_0^1 \frac{dt}{1 + t + \dots + t^n}.$$

54.2 Montrer, en utilisant une série géométrique et l'égalité $\sum_{n=0}^{+\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$, que :

$$\int_0^{+\infty} \frac{t}{\text{sh } t} dt = \frac{\pi^2}{4}.$$

54.3 Soit $x > -1$. Définition et continuité de : $f(x) = \int_0^1 \frac{\ln(1 + xt^2)}{t^2} dt$.

54.4 Soit $x \in]0; \pi[$. Continuité et dérivabilité de : $f(x) = \int_0^1 \frac{\ln(t^2 - 2t \cos x + 1)}{t} dt$.

54.5

www

Calculer, pour tout $t \in \mathbb{R}$, $\int_{\mathbb{R}} e^{-x^2 - itx} dx$.

54.6 Soient $f(x, t) = \frac{1 - \cos t}{\sqrt{\sin^2 t + x \cos^2 t}}$ et $F(x) = \int_0^{\frac{\pi}{2}} f(x, t) dt, \forall (t, x) \in]0; \frac{\pi}{2}[\times \mathbb{R}_+$.

1. Montrer que F est décroissante sur $\mathbb{R}_+$ et continue sur $\mathbb{R}_+^*$.

2. En utilisant le théorème de convergence monotone, montrer que F est continue en 0 et calculer $\lim_{x \rightarrow +\infty} F(x)$.

3. En appliquant le théorème de Leibniz, montrer que F est de classe $\mathcal{C}^1$ sur $\mathbb{R}_+^*$.

Équations différentielles : premier ordre

1. Généralités

Définition 55.1

Une **équation différentielle** est une relation entre une variable réelle t et les valeurs $x(t)$, $x'(t)$, $\dots$, $x^{(n)}(t)$ d'une fonction inconnue et de certaines de ses dérivées, de la forme :

$$F(t, x(t), x'(t), \dots, x^{(n)}(t)) = 0 \quad (E)$$

où F est une fonction continue.

L'**ordre** n de l'équation différentielle est celui de la dérivée d'ordre le plus élevé qui figure dans (E) .

Une **solution** de (E) sur I , ou intégrale de (E) sur I , est une fonction définie sur l'intervalle ouvert I possédant des dérivées jusqu'à l'ordre n et vérifiant (E) .



La variable est parfois notée x et la fonction inconnue notée $y(x)$. Dans la section 2, c'est cette notation que nous prenons.

Définition 55.2 Équations différentielles linéaires

Soient a , b et f des fonctions données, continues sur un intervalle I , à valeurs réelles et complexes. Une **équation différentielle linéaire d'ordre 1** est de la forme :

$$\forall t \in I, \quad a(t)x'(t) + b(t)x(t) = f(t). \quad (55.1)$$

Pour la résoudre, il faut se placer sur un intervalle $J \subset I$ sur lequel a ne s'annule pas. L'équation est dite à coefficients constants si a et b sont constants avec $a \neq 0$.

Théorème 55.1

• Toute solution de (55.1) est de la forme $x_P(t) + x_S(t)$ où $x_P(t)$ est une solution particulière de (55.1) et $x_S(t)$ la solution générale de l'équation homogène associée :

$$a(t)x'(t) + b(t)x(t) = 0. \quad (55.2)$$

- Les solutions complexes de (55.2) sur J forment un $\mathbb{C}$ -espace vectoriel de dimension 1.
- Si a et b sont à valeurs réelles, les solutions réelles de (55.2) sur J forment un $\mathbb{R}$ -espace vectoriel de dimension 1.

Théorème 55.2

Les solutions de l'équation homogène associée sont du type :

$$x_S(t) = Ke^{-A(t)} \quad \text{où} \quad A(t) = \int_{t_0}^t \frac{b(u)}{a(u)} du$$

avec K une constante réelle ou complexe et t_0 un élément quelconque de J .

Théorème 55.3 Variation de la constante

x_1 étant une solution non nulle de (55.2), nous introduisons une fonction auxiliaire inconnue $K(t)$ pour $t \in J$ telle que $x(t) = K(t)x_1(t)$ soit solution de (55.1). Par conséquent, $K'(t) = \frac{f(t)}{a(t)x_1(t)}$ ce qui permet de calculer $K(t)$ puis $x(t)$.

2. Situations particulières

Définition 55.3 Équations à variables séparées

Soient f une fonction continue sur un intervalle I de $\mathbb{R}$ et g une fonction continue et non nulle sur un intervalle J de $\mathbb{R}$, une équation à variables séparées est une équation différentielle du premier ordre pouvant s'écrire sous la forme : $g(y)y' = f(x)$, i.e. $\forall x \in I, g(y(x)) \times y'(x) = f(x)$.

Théorème 55.4

Toute solution d'une équation à variables séparées $g(y)y' = f(x)$ est définie par $G(y) = F(x) + \lambda$, G et F sont des primitives de g et f sur J et sur I avec $\lambda \in \mathbb{R}$.

Définition 55.4 Équations différentielles homogènes

Nous appelons équation différentielle homogène du premier ordre une équation de la forme : $f(x, y, y') = 0$ dans laquelle le changement x en λx et y en λy laisse y' invariant. Nous pouvons écrire aussi : $F\left(\frac{y}{x}, y'\right) = 0$.

Théorème 55.5

Pour résoudre une équation différentielle homogène du premier ordre nous cherchons d'abord les **solutions singulières** qui sont les fonctions linéaires $x \mapsto mx$ où m est une solution de $F(m, m) = 0$. Nous cherchons ensuite la solution générale :

- soit en utilisant une paramétrisation de la courbe d'équation : $F(X, Y) = 0$;
- soit en utilisant les coordonnées polaires : $x = r \cos(t)$ et $y = r \sin(t)$.



Si nous avons la forme : $f(x, y') = 0$ ou $f(y, y') = 0$, nous utiliserons une paramétrisation de la courbe $f(X, Y) = 0$.

Théorème 55.6 Équation de Bernoulli

Une équation différentielle de Bernoulli est une équation de la forme :

$$a(x)y'(x) + b(x)y(x) = f(x)y^\alpha(x)$$

où a , b et c sont des fonctions continues sur $\mathbb{R}$ et α est un réel.

Si $\alpha = 1$, l'équation est linéaire. Si $\alpha \neq 1$, nous posons $z = y^{1-\alpha}$ et la résolution de l'équation se ramène à la résolution d'une équation linéaire en z et z' .

Théorème 55.7 Équation de Riccati

Une équation différentielle de Riccati peut s'écrire sous la forme :

$$y' = a(x)y^2 + b(x)y + c(x)$$

où a , b et c sont des fonctions continues sur $\mathbb{R}$. Pour la résoudre, il suffit de connaître une solution particulière φ de l'équation puis nous cherchons la solution générale sous la forme $y = \varphi + \frac{1}{u}$.

Application

1. Chute libre freinage

Un corps de masse m en chute libre subit une force de freinage : $\vec{F} = -k\vec{v}$. La relation fondamentale de la dynamique donne : $\frac{dv}{dt} + \frac{k}{m}v = g$ où $v = \|\vec{v}\|$, k une constante non nulle et $g = 9,81$. Résoudre cette équation différentielle.

2. Équation différentielle homogène

Résoudre $xy' - y(1 + y^2) = 0$.

3. Équation différentielle à second membre

Résoudre l'équation différentielle : $y' + 2y = 2 \operatorname{sh}(2x)$.

4. Équations de Bernoulli

Résoudre les équations différentielles :

$$\text{a. } y' + 2y - (x+1)\sqrt{y} = 0. \quad \text{b. } xy' + y = \frac{y^2}{x^2} \text{ avec } y(1) = 1.$$

1. La solution de l'équation différentielle sans second membre est égale à $v_S(t) = Ae^{-\frac{kt}{m}}$, où $A \in \mathbb{R}$.

Une solution particulière de l'équation différentielle avec second membre est égale à $v_P(t) = \frac{mg}{k}$.

La solution générale de cette équation différentielle linéaire est égale à $v(t) = Ae^{-\frac{kt}{m}} + \frac{mg}{k}$.

À l'instant initial, $v(0) = 0$ donc $A = \frac{-mg}{k}$. Finalement nous avons $v(t) = \frac{mg}{k} \left(1 - e^{-\frac{kt}{m}}\right)$.

2. C'est une équation différentielle homogène. Nous pouvons l'écrire, si $x \neq 0$, $y' - \frac{y}{x}(1 + y^2) = 0$.

Soit $y(x) = mx$ une solution homogène. Nous avons alors $mx - mx(1 + m^2) = 0$, pour tout $x \in \mathbb{R}$. Ainsi $m = 0$ et par conséquent $y(x) = 0$ est une solution singulière.

Nous avons $F(X, Y) = Y - X(1 + Y^2) = 0$. Donc, en utilisant la paramétrisation suivante $X = \frac{t}{1+t^2}$ et $Y = t$, nous avons $y = xX = \frac{tx}{1+t^2}$ et $y' = t$. Nous avons alors

$dy = \frac{tdx}{1+t^2} + \frac{x(1-t^2)dt}{(1+t^2)^2}$ et $dy = tdx$. Donc, nous pouvons écrire $\frac{dx}{x} = \frac{(1-t^2)dt}{(1+t^2)t^3}$. Nous

posons comme changement de variable $u = \frac{1}{t^2}$ ce qui implique que $du = \frac{-2dt}{t^3}$. Nous avons

alors $\frac{dx}{x} = \frac{(1-u)du}{2(1+u)}$. En intégrant, nous obtenons $\ln\left(\frac{x}{k}\right) = \ln\left(e^{\frac{-u}{2}}\right) + \ln|1+u|$. En conclu-

sion, nous avons $x = k\left(1 + \frac{1}{t^2}\right)e^{\frac{-1}{2t^2}}$ et $y = \frac{k}{t}e^{\frac{-1}{2t^2}}$ où $k \in \mathbb{R}^*$.

3. L'équation différentielle sans second membre a pour solution $v(x) = ke^{-2x}$, $k \in \mathbb{R}$.

Nous allons maintenant faire varier la constante k , nous posons alors $v(x) = k(x)e^{-2x}$.

$v(x)$ est solution de $y' + 2y = e^{2x} - e^{-2x}$ si et seulement si $e^{-2x}k'(x) = e^{2x} - e^{-2x}$ ou encore $k'(x) = e^{4x} - 1$. En intégrant, $k(x) = \frac{1}{4}e^{4x} - x + \alpha$ où $\alpha \in \mathbb{R}$. Donc nous avons $v(x) = \alpha e^{-2x} + \frac{1}{4}e^{2x} - xe^{-2x}$ où $\alpha \in \mathbb{R}$.

4. a. C'est une équation de Bernoulli qui admet la solution singulière $y = 0$.

En simplifiant par $\sqrt{y}$, l'équation initiale se transforme en $\frac{y'}{\sqrt{y}} + 2\sqrt{y} = x + 1$. Nous posons le changement de variable $z = \sqrt{y}$, ce qui implique que $z' = \frac{y'}{2\sqrt{y}}$.

L'équation se transforme ainsi : $2z' + 2z = x + 1$ qui est une équation différentielle linéaire du premier ordre.

La solution générale de l'équation homogène associée à $z' + z = 0$ est égale à : $z_1(x) = Ae^{-x}$, où A est une constante réelle.

La solution $z_2(x) = \frac{x}{2}$ est une solution particulière de (a). Ainsi $z(x) = Ae^{-x} + \frac{x}{2}$, où A est une constante réelle. Finalement, nous avons $y(x) = \left(Ae^{-x} + \frac{x}{2}\right)^2$, où A est une constante réelle.

b. Nous posons le changement de variable $z = y^{1-2} = y^{-1}$, ce qui implique $z' = \frac{-y'}{y^2}$. En

divisant par x , l'équation différentielle s'écrit sous la forme : $y' + \frac{y}{x} = \frac{y^2}{x^{\frac{5}{2}}}$. Nous avons

également $y' = \frac{-z'}{z^2}$ et par conséquent l'équation différentielle se transforme en $\frac{-z'}{z^2} + \frac{1}{zx} = \frac{1}{z^2 x^{\frac{5}{2}}}$. En multipliant cette dernière par $-1/z^2$, nous obtenons : $z' - \frac{z}{x} = \frac{-1}{x^{\frac{5}{2}}}$.

Ainsi, $\left(\frac{z}{x}\right)' = \frac{-1}{x^{\frac{7}{2}}}$. En intégrant, et en prenant en compte que $y(1) = 1$, nous obtenons $\frac{z}{x} = \frac{2}{5} \frac{1}{x^{\frac{5}{2}}} + \frac{3}{5}$. Par conséquent nous avons $\frac{1}{y} = \frac{2}{5} \frac{1}{x^{\frac{3}{2}}} + \frac{3x}{5}$. Finalement $y = \frac{5x^{\frac{3}{2}}}{2 + 3x^{\frac{5}{2}}}$.

Pour s'entraîner (solutions p. 532)

55.1 Résoudre les équations différentielles :

1. $y' - y \tan(x) = \frac{1}{1 + \cos(x)}$, $x \in \left]-\frac{\pi}{2}; \frac{\pi}{2}\right[$.

2. $xy' - 4y^2(1 - x) = 0$.

3. $y' \sqrt{x} - y + \sqrt{y}(x - 2\sqrt{x}) = 0$.

55.2 

1. Soit la fonction ϕ définie sur $\mathbb{R}_+$ par :

pour $x = 0$, $\phi(0) = 1$

pour $x > 0$, $\phi(x) = x/(e^x - 1)$.

a. Montrer que ϕ est de classe $\mathcal{C}^1$ sur $\mathbb{R}_+$.

b. Établir le tableau de variation de ϕ .

c. Établir le développement limité d'ordre 2 de ϕ en 0.

2. Soit l'équation différentielle

$$y'(1 - e^{-x}) + y = e^{-x}. \quad (55.3)$$

a. Montrer que ϕ est une solution de (55.3) sur $\mathbb{R}_+$, puis résoudre l'équation (55.3) sur $\mathbb{R}_+$.

b. Montrer qu'il n'existe pas d'autre solution de (55.3) avec une limite finie en 0.

c. ϕ est-elle aussi une solution de (55.3) sur $\mathbb{R}_+$?

Équations différentielles : second ordre

Définition 56.1 Équations différentielles linéaires

Soient a, b, c et f des fonctions données, continues sur un intervalle I , à valeurs réelles et complexes. Une **équation différentielle linéaire d'ordre 2** est de la forme :

$$a(t)x''(t) + b(t)x'(t) + c(t)x(t) = f(t). \quad (56.1)$$

Pour la résoudre, il faut se placer sur un intervalle $J \subset I$ sur lequel a ne s'annule pas. L'équation est dite à coefficients constants si a, b et c sont constants avec $a \neq 0$.



La variable est parfois notée x et la fonction inconnue notée $y(x)$.

Théorème 56.1

- Toute solution de (56.1) est de la forme $x_p(t) + x_s(t)$ où $x_p(t)$ est une solution particulière de (56.1) et $x_s(t)$ la solution générale de l'équation homogène associée :

$$a(t)x''(t) + b(t)x'(t) + c(t)x(t) = 0. \quad (56.2)$$

- Les solutions complexes de (56.2) sur J forment un $\mathbb{C}$ -espace vectoriel de dimension 2.
- Si a, b et c sont à valeurs réelles, les solutions réelles de (56.2) sur J forment un $\mathbb{R}$ -espace vectoriel de dimension 2.

Théorème 56.2 Cas homogène à coefficients constants

Dans le cas où les coefficients a, b, c dans l'équation différentielle (56.2) sont constants, avec $a \neq 0$, l'équation $ar^2 + br + c = 0$ s'appelle **équation caractéristique** de l'équation différentielle (56.2). Notons $\Delta = b^2 - 4ac$, son discriminant.

- Si $\Delta \neq 0$, l'équation caractéristique a deux racines réelles distinctes r_1 et r_2 et :

$$x_s(t) = K_1 \exp(r_1 t) + K_2 \exp(r_2 t),$$

où K_1 et K_2 sont des constantes réelles quelconques.

- Si $\Delta = 0$, l'équation caractéristique a une racine réelle double r_0 et :

$$x_s(t) = (K_1 t + K_2) \exp(r_0 t),$$

où K_1 et K_2 sont des constantes réelles quelconques.

- Si a et b sont deux réels et si $\Delta < 0$, l'équation caractéristique a deux racines complexes conjuguées $\alpha \pm i\beta$:

$$x_s(t) = \exp(\alpha t)(K_1 \cos(\beta t) + K_2 \sin(\beta t)),$$

où K_1 et K_2 sont des constantes réelles quelconques.

Théorème 56.3 Solutions particulières à coefficients constants

Cas où $f(t)$ est un polynôme $P(t)$ de degré n

Il existe une solution particulière de (56.1) sous la forme d'un polynôme de degré :

- n si $c \neq 0$;
- $n + 1$ si $c = 0$ et $b \neq 0$;
- $n + 2$ si $c = 0$ et $b = 0$.

La recherche de cette solution se fait par identification.

Cas où $f(t) = \exp(kt)P(t)$ avec P polynôme et k une constante réelle ou complexe

Procéder au changement de fonction inconnue $x(t) = \exp(kt)z(t)$ et se ramener au cas précédent.

Cas où $f(t) = \cos(k_1t)P(t) + \sin(k_2t)Q(t)$ avec P et Q deux polynômes et k_1 et k_2 deux constantes réelles

Il est possible de se ramener au cas précédent en constatant que $\cos(k_1t)P(t)$ est la partie réelle de $\exp(ik_1t)P(t)$ et $\sin(k_2t)Q(t)$ est la partie imaginaire de $\exp(ik_2t)Q(t)$. La solution particulière recherchée est la somme des deux solutions particulières obtenues précédemment.



Soit (E) : $ax'' + bx' + cx = \exp(\alpha t) \cos(mt) P(t)$ ou encore $\exp(\alpha t) \sin(mt) P(t)$. Nous nous ramenons au cas précédent en utilisant les formules d'Euler $\cos(mt) = \frac{e^{imt} + e^{-imt}}{2}$ et $\sin(mt) = \frac{e^{imt} - e^{-imt}}{2i}$.

Théorème 56.4 Variation de la constante

Nous cherchons une solution particulière de (56.1) de la forme : $x(t) = \lambda_1(t)x_1(t) + \lambda_2(t)x_2(t)$, où x_1 et x_2 sont deux solutions linéaires indépendantes de l'équation sans second membre, ce qui implique que, pour tout $t \in I$, $x_1(t)x_2'(t) - x_2(t)x_1'(t) \neq 0$. Pour tout $t \in I$, les fonctions λ_1, λ_2 sont définies comme les uniques solutions du système :

$$\begin{cases} \lambda_1'(t)x_1(t) + \lambda_2'(t)x_2(t) = 0 \\ \lambda_1'(t)x_1'(t) + \lambda_2'(t)x_2'(t) = \frac{f(t)}{a(t)}. \end{cases}$$

Définition 56.2 Second ordre et premier ordre

L'intégration de certaines équations du second ordre peut se ramener à celle d'équations du premier ordre.

Soit une équation du second ordre de la forme : $F(t, x', x'') = 0$, cette équation ne contient pas x . Si nous posons $x' = z$, l'équation devient : $F(t, z, z') = 0$, z est donc solution d'une équation du premier ordre.

Soit une équation du second ordre de la forme : $F(x, x', x'') = 0$, cette équation ne contient pas de t . Si nous posons $x' = z(x)$, nous avons $x'' = \frac{dx'}{dt} = \frac{dz}{dx} \frac{dx}{dt} = \frac{dz}{dx} z$, donc, nous obtenons : $F\left(x, z, z \frac{dz}{dx}\right) = 0$.

z est donc solution d'une équation du premier ordre.

1. Coefficients constants et homogène

Résoudre l'équation différentielle : $y'' - 3y' + 2y = 0$.

2. Coefficients constants avec second membre particulier

Résoudre l'équation différentielle : $y'' - 2y' + 2y = x \cos(x)$ (E).

3. Coefficients constants avec second membre

Résoudre l'équation différentielle : $y'' + y = \tan(x)$.

4. Équation différentielle avec second membre

Résoudre l'équation différentielle : $x^2 y'' + xy' - 4y = x^2 \ln(x)$ (1).

5. Équation différentielle non linéaire

Résoudre l'équation différentielle : $y'' + y'^2 = 0$.

1. L'équation caractéristique est $r^2 - 3r + 2 = 0$ avec $\Delta = 1$. Les racines de cette équation sont : $r_1 = 1$ et $r_2 = 2$. La solution générale de l'équation différentielle est : $y(x) = K_1 \exp(x) + K_2 \exp(2x)$, avec $(K_1, K_2) \in \mathbb{R}^2$.

2. L'équation caractéristique de l'équation homogène (H) associée à (E) est égale à $r^2 - 2r + 2 = 0$ avec $\Delta = -4$. Ainsi $r_1 = 1 + i$ et $r_2 = 1 - i$. Donc, les solutions de (H) sont égales à $y_1(x) = \exp(x) (K_1 \cos(x) + K_2 \sin(x))$, où $(K_1, K_2) \in \mathbb{R}^2$. Utilisons le fait que : $x \cos(x) = \frac{x}{2} (e^{ix} + e^{-ix})$.

Une solution particulière y_2 de l'équation $y'' - 2y' + 2y = \frac{x}{2} e^{ix}$ est $y_2(x) = \left(\frac{1+2i}{10} x + \frac{1+7i}{25} \right) e^{ix}$.

Une solution particulière y_3 de l'équation $y'' - 2y' + 2y = \frac{x}{2} e^{-ix}$ est $y_3(x) = \left(\frac{1-2i}{10} x + \frac{1-7i}{25} \right) e^{-ix}$.

Donc une solution particulière de (E), qui s'obtient en additionnant y_1 et y_2 , est égale à $y(x) = \left(\frac{1+2i}{10} x + \frac{1+7i}{25} \right) e^{ix} + \left(\frac{1-2i}{10} x + \frac{1-7i}{25} \right) e^{-ix} = \frac{5x+2}{25} \cos(x) - \frac{10x+14}{25} \sin(x)$.

Les solutions générales de (E) sont égales à $y(x) = \exp(x) (K_1 \cos(x) + K_2 \sin(x)) + \frac{5x+2}{25} \cos(x) - \frac{10x+14}{25} \sin(x)$ pour tout $x \in \mathbb{R}$ avec $(K_1, K_2) \in \mathbb{R}^2$.

3. Une solution de $y'' + y = 0$ est $y = k_1 \cos(x) + k_2 \sin(x)$, où $(k_1, k_2) \in \mathbb{R}^2$.

Maintenant appliquons la méthode de la variation de la constante. Nous avons $y' = -k_1(x) \sin(x) + k_2(x) \cos(x) + k_1'(x) \cos(x) + k_2'(x) \sin(x)$. Nous imposons $k_1'(x) \cos(x) + k_2'(x) \sin(x) = 0$. Ainsi nous obtenons $y' = -k_1(x) \sin(x) + k_2(x) \cos(x)$. Par conséquent nous avons $y'' = -k_1'(x) \sin(x) + k_2'(x) \cos(x) - k_1(x) \cos(x) - k_2(x) \sin(x)$. Donc, en reportant dans l'équation différentielle, nous avons :

$$\begin{cases} k_1'(x) \cos(x) + k_2'(x) \sin(x) = 0 \\ -k_1'(x) \sin(x) + k_2'(x) \cos(x) = \tan(x) \end{cases}$$

Donc $k_1'(x) = \cos(x) - \frac{1}{\cos(x)}$ et $k_2'(x) = \sin(x)$. Puis nous intégrons et nous obtenons $k_1(x) = \sin(x) - \ln \left| \tan \left(\frac{x}{2} + \frac{\pi}{4} \right) \right| + C_1$ et $k_2(x) = -\cos(x) + C_2$ où $(C_1, C_2) \in \mathbb{R}^2$. La solution générale est égale à $y(x) = C_1 \cos(x) + C_2 \sin(x) - \cos(x) \ln \left| \tan \left(\frac{x}{2} + \frac{\pi}{4} \right) \right|$ où $(C_1, C_2) \in \mathbb{R}^2$.

4. Une solution particulière de $x^2 y'' + xy' - 4y = 0$ (2) est $y = x^2$.

Nous posons $y = y_1 z$ dans (2), nous avons alors $x^2(y_1''z + 2y_1'z' + y_1z'') + x(y_1'z + y_1z') - 4y_1z = 0$.

Donc, $x^4 z'' + 5x^3 z' = 0$ soit avec $u = z'$, $x^4 u' + 5x^3 u = 0$. Pour $u \neq 0$, $\frac{u'}{u} = \frac{-5}{x} \Rightarrow \ln |u| = -5 \ln(x) + k_1$, $k_1 \in \mathbb{R} \Rightarrow u = \frac{-k_1}{4x^4} = \frac{k_2}{x^4}$, $k_2 \in \mathbb{R}$, donc $y_1 z = y_2 = \frac{k_2}{x^2}$, la solution générale de (2) est alors $y = kx^2 + \frac{k_2}{x^2}$.

Nous posons maintenant $y = \lambda_1(x)y_1 + \lambda_2(x)y_2$ en imposant $\lambda_1' y_1 + \lambda_2' y_2 = 0$ soit $\lambda_1' x^2 + \frac{\lambda_2'}{x^2} = 0$.

Par conséquent, $y' = 2\lambda_1 x - 2\frac{\lambda_2}{x^3}$ et $y'' = 2\lambda_1' x + 2\lambda_1 - \frac{2\lambda_2'}{x^3} + \frac{6\lambda_2}{x^4}$, ce qui donne dans (1)

$2\lambda_1' x^3 - \frac{2\lambda_2'}{x} = x^2 \ln(x)$. Nous obtenons alors $\lambda_2' = \frac{-x^3 \ln(x)}{4}$ donc $\lambda_2(x) = \frac{-x^4 \ln(x)}{16} + \frac{x^4}{64}$ et

$$\lambda_1(x) = \frac{\ln^2(x)}{8}.$$

La solution générale est donc :

$$y = k_1 x^2 + \frac{k_2}{x^2} + \frac{x^2 \ln^2(x)}{8} - \frac{x^2 \ln(x)}{16} + \frac{x^2}{64} \quad \forall (k_1, k_2) \in \mathbb{R}^2.$$

5. Nous posons $z = y'$, nous avons alors : $z' + z^2 = 0$ donc $\frac{dz}{dx} + z^2 = 0 \iff \frac{-dz}{z^2} = dx \iff \frac{1}{z} = x + k_1$ avec k_1 : constante.

Donc, $z = \frac{dy}{dx} = \frac{1}{x + k_1}$, nous obtenons donc :

$$y + k_2 = \ln |x + k_1| \quad \text{avec } k_2 \text{ une constante.}$$

Pour s'entraîner (solutions p. 532)

56.1 Résoudre les équations différentielles :

1. $y'' + y = \cos(x)$
2. $y'' + y' - 2y = 8 \sin(2x)$
3. $y'' + y' - 2y = xe^{-2x}$
4. $y'' - 2y' + y = 16 \operatorname{ch}(x) \cos(x)$
5. $y'' + 2y' - 8y = 5e^{2x}(3x + 9)$
6. $y'' - 4y' + 13y = 12 \cos(2x) + 28 \sin(2x)$

56.2 

Notons λ un paramètre réel fixé et définissons la suite $(a_n)_{n \in \mathbb{N}}$ par $a_0 = 1$, et

$$\text{pour } n \geq 1, a_n = \frac{\prod_{k=0}^{n-1} (\lambda + k)}{(n!)^2}.$$

1. Déterminer le rayon de convergence de la série entière $\sum a_n x^n$.

Dans la suite, nous définissons pour $x \in]-R; R[$, la fonction $f_\lambda(x) = \sum_{n=0}^{+\infty} a_n x^n$.

2. Exprimer f_{-2} , f_{-1} , f_0 et f_1 à l'aide de fonctions usuelles.

3. Montrer que si $p \in \mathbb{N}$, f_{-p} est une fonction polynomiale, dont vous préciserez le degré et le terme de plus haut degré.

4. Montrer qu'il existe une unique fonction $x \mapsto S(x)$ développable en série entière, de rayon de convergence non nul, solution du problème

$$\begin{cases} xy''(x) + (1-x)y'(x) - \lambda y(x) = 0 \\ y(0) = 1 \end{cases} \quad (E_\lambda)$$

Exprimer cette solution en fonction de f_λ .

5. Montrer que la fonction

$$g_\lambda : x \mapsto e^x f_\lambda(-x)$$

est développable en série entière.

Donner son rayon de convergence.

6. Montrer que g_λ est solution d'une équation différentielle dont vous donnerez une expression. Dédurre une relation entre g_λ et $f_{1-\lambda}$.

Fonction d'une variable réelle : Exercices de synthèse

Application

Résoudre l'équation différentielle : $y^2 y'' + y' = 0$.

Nous posons $y' = z(y)$ donc $y'' = zz'$, nous obtenons $y^2 zz' + z = 0$.

$z = 0$ est solution, nous n'en tenons pas compte, donc : $y^2 z' + 1 = 0 \iff z' = \frac{-1}{y^2} \iff z = \frac{1}{y} + k_1 = \frac{1 + k_1 y}{y} = \frac{dy}{dx}$ où k_1 est une constante.

Nous obtenons : $dx = \frac{y dy}{1 + k_1 y} = \frac{1}{k_1} \left(1 - \frac{1}{1 + k_1 y} \right) dy$.

Donc, $x = \frac{y}{k_1} - \frac{1}{k_1^2} \ln |1 + k_1 y| + k_2$, où k_2 est une constante.

Application Circuit RLC

Dans le cas d'un circuit RLC en régime libre, lorsque nous fermons le circuit et que le condensateur se décharge dans ce circuit, la tension $u(t)$ aux bornes du condensateur est régie par une équation différentielle homogène à coefficients constants du deuxième ordre :

$$u''(t) + 2mw_0 u'(t) + w_0^2 u(t) = 0 \quad (E), \text{ avec } w_0 = \frac{1}{\sqrt{LC}} \text{ et } m = \frac{R}{2} \sqrt{\frac{C}{L}} > 0.$$

Résoudre (E) dans les trois cas suivants : $m > 1$, $m = 1$, $m < 1$.

L'équation caractéristique de l'équation différentielle est $r^2 + 2mw_0 r + w_0^2 = 0$ avec $\Delta = 4w_0^2(m^2 - 1)$.

Si $m > 1$, $\Delta > 0$, ce qui implique deux racines réelles : $r_1 = -mw_0 + w_0 \sqrt{m^2 - 1}$ et $r_2 = -mw_0 - w_0 \sqrt{m^2 - 1}$ donc $u(t) = \alpha e^{r_1 t} + \beta e^{r_2 t}$. Si le condensateur est chargé à $t = 0s$, nous avons $u(0) = E$ et le courant est nul : $i(0) = 0$, or, $i(t) = C \frac{du(t)}{dt} = (\alpha r_1 e^{r_1 t} + \beta r_2 e^{r_2 t}) C$. Nous

avons alors $\alpha + \beta = E$ et $\alpha r_1 + \beta r_2 = 0$ donc $\beta = \frac{E(\sqrt{m^2 - 1} - m)}{2\sqrt{m^2 - 1}}$ et $\alpha = \frac{E(\sqrt{m^2 - 1} + m)}{2m}$.

En conclusion :

$$u(t) = \frac{E(\sqrt{m^2 - 1} + m)}{2m} e^{(-mw_0 + w_0 \sqrt{m^2 - 1})t} + \frac{E(\sqrt{m^2 - 1} - m)}{2\sqrt{m^2 - 1}} e^{-(mw_0 + w_0 \sqrt{m^2 - 1})t}.$$

Si $m = 1$, $\Delta = 0$ donc l'équation caractéristique a une racine double : $r = -w_0$, donc $u(t) = (\alpha t + \beta) e^{-w_0 t}$ or $u(0) = E$ et $i(0) = 0$ donc $\alpha = E w_0$. Nous avons alors :

$$u(t) = E (w_0 t + 1) e^{-w_0 t}.$$

Si $m < 1$, $\Delta < 0$ donc l'équation caractéristique a deux racines conjuguées : $r_1 = -mw_0 + jw_0 \sqrt{1-m^2}$ et $r_2 = -mw_0 - jw_0 \sqrt{1-m^2}$ donc, $u(t) = e^{-mw_0 t} (\alpha \cos(w_0 \sqrt{1-m^2} t) + \beta \sin(w_0 \sqrt{1-m^2} t))$. Or, $u(0) = \alpha = E$ et $i(0) = 0$.
 $i(t) = -C m w_0 e^{-mw_0 t} (\alpha \cos(w_0 \sqrt{1-m^2} t) + \beta \sin(w_0 \sqrt{1-m^2} t))$
 $+ C e^{-mw_0 t} (-\alpha w_0 \sqrt{1-m^2} \sin(w_0 \sqrt{1-m^2} t) + \beta w_0 \sqrt{1-m^2} \cos(w_0 \sqrt{1-m^2} t))$. Donc, $\beta = \frac{mE}{\sqrt{1-m^2}}$. Puis, $u(t) = \frac{E}{\sqrt{1-m^2}} e^{-mw_0 t} (\sqrt{1-m^2} \cos(w_0 \sqrt{1-m^2} t) + m \sin(w_0 \sqrt{1-m^2} t))$.
 Nous posons : $\cos(\varphi) = \sqrt{1-m^2}$, $\sin(\varphi) = m$, donc $u(t) = \frac{E}{\sqrt{1-m^2}} e^{-mw_0 t} (\cos(\varphi) \cos(w_0 \sqrt{1-m^2} t) + \sin(\varphi) \sin(w_0 \sqrt{1-m^2} t))$. Nous avons alors :

$$u(t) = \frac{E}{\sqrt{1-m^2}} e^{-mw_0 t} (\cos(w_0 \sqrt{1-m^2} t - \varphi)) \text{ avec } \tan(\varphi) = \frac{m}{\sqrt{1-m^2}}.$$

Application Équation de Bessel série entière

Nous appelons équation de Bessel l'équation différentielle du second ordre :

$$y'' + \frac{1}{x} y' + \left(1 - \frac{k^2}{x^2}\right) y = 0, \text{ où } k \text{ est une constante.}$$

Cette équation se rencontre dans un grand nombre de problèmes de mécanique et d'électricité. Nous cherchons les solutions sous la forme d'une série entière : $y = x^m \sum_{n=0}^{+\infty} a_n x^n$.

Résoudre alors l'équation différentielle.

Nous avons :

$$\frac{-k^2}{x^2} y = -k^2 (a_0 x^{m-2} + a_1 x^{m-1} + \dots + a_n x^{m+n-2} + a_{n+1} x^{m+n-1} + a_{n+2} x^{m+n} + \dots)$$

$$\frac{y'}{x} = a_0 m x^{m-2} + a_1 (m+1) x^{m-1} + \dots + a_n (m+n) x^{m+n-2} + a_{n+1} (m+n+1) x^{m+n-1} + a_{n+2} (m+n+2) x^{m+n} + \dots$$

$$y'' = a_0 m(m-1) x^{m-2} + a_1 (m+1) m x^{m-1} + \dots + a_n (m+n)(m+n-1) x^{m+n-2} + a_{n+1} (m+n+1)(m+n) x^{m+n-1} + a_{n+2} (m+n+2)(m+n+1) x^{m+n} + \dots$$

$$\text{Donc, } a_{n+2} = \frac{a_n}{k^2 - (m+n+2)^2} \text{ et } a_0 (m^2 - k^2) = 0, a_1 = 0 \text{ donc } m = \pm k.$$

$$\text{Si } m = k, a_{n+2} = \frac{-a_n}{(2k+n+2)(n+2)} \Rightarrow a_n = \frac{-a_{n-2}}{(2k+n)n}.$$

$$\text{Si } n = 2p+1, a_{2p+1} = 0 \text{ car } a_1 = 0.$$

$$\text{Si } n = 2p, a_{2p} = \frac{-a_{2p-2}}{2p(2k+2p)}, \dots, a_2 = \frac{-a_0}{2(2k+2)}. \text{ Donc}$$

$$a_{2p} = \frac{(-1)^p}{2^{2p}} \frac{a_0}{p!(k+1)(k+2) \dots (k+p)}.$$

Donc

$$y_k = x^k \sum_{p=0}^{+\infty} (-1)^p \frac{a_0}{p!(k+1)(k+2)\cdots(k+p)} \left(\frac{x}{2}\right)^{2p}.$$

Si $m = -k$, nous avons alors :

$$y_{-k} = x^{-k} \sum_{p=0}^{+\infty} (-1)^p \frac{a_0}{p!(1-k)(2-k)\cdots(p-k)} \left(\frac{x}{2}\right)^{2p}.$$

y_k et y_{-k} étant linéairement indépendantes, nous avons : $y = k_1 y_k + k_2 y_{-k}$, où k_1, k_2 sont des constantes.

Si $k = n > 0$ et $a_0 = \frac{1}{2^n n!}$, nous obtenons la fonction de Bessel d'ordre n :

$$J_n(x) = \left(\frac{x}{2}\right)^n \sum_{p=0}^{+\infty} \frac{(-1)^p}{p!(n+p)!} \left(\frac{x}{2}\right)^{2p}.$$

Pour s'entraîner (solutions p. 532)

57.1 Résoudre l'équation différentielle : $y^2 y'' - 2y y'^2 - y'^3 = 0$.

57.2 Résoudre l'équation différentielle : $yy'' - 2y'^2 + 4y^2 = 0$.

57.3 Résoudre l'équation différentielle : $xy'' + 2y' + xy = 0$ en vérifiant que la fonction $f(x) = \frac{\sin(x)}{x}$ est solution de l'équation différentielle.

57.4 Résoudre l'équation différentielle : $xy'' - y' - x^3 y = 0$.

57.5 Résoudre l'équation différentielle : $(x^2 - 3)y'' - 4xy' + 6y = 0$.

57.6 Résoudre l'équation différentielle : $x^2 y'' - 3xy' + 4y = x^3$.

57.7  [www](#)

1. Montrer l'existence de l'intégrale :

$$I = \int_0^{+\infty} \frac{x}{e^x - 1} dx$$

et exprimer I à l'aide d'une série.

2. Trouver a et b réels tels que, pour tous $n \in \mathbb{N}^*$:

$$\int_0^\pi (at + bt^2) \cos(nt) dt = \frac{1}{n^2}.$$

En déduire la valeur de I .

57.8 Soit $f : x \mapsto \sum_{n=1}^{+\infty} \frac{\arctan(nx)}{n^2}$.

1. Étudier les variations de f .
2. Calculer la limite de f en $+\infty$.
3. Préciser un équivalent de f lorsque $x \rightarrow 0$.

57.9 Considérons la série de fonctions $x \mapsto \sum_{n=1}^{+\infty} \frac{1}{\operatorname{sh}(nx)}$.

1. Trouver un équivalent de f en 0.
2. Préciser un équivalent de f lorsque $x \rightarrow +\infty$.

57.10 Soit $f : x \mapsto \sum_{n=1}^{+\infty} e^{-x\sqrt{n}}$. Calculer un équivalent de f lorsque $x \rightarrow 0$.

57.11 Soit f la fonction définie par :

$$f(x) = \int_0^{+\infty} e^{-xt} \frac{\sin(t)}{t} dt.$$

1. Montrer que f est de classe $\mathcal{C}^1$ sur $[0; +\infty[$.
2. Exprimer f' puis f sans avoir recours à une intégrale.

57.12 Soit F la fonction définie par :

$$F(x) = \int_0^{+\infty} \frac{\ln(x^2 + t^2)}{1 + t^2} dt.$$

1. Donner l'ensemble de définition de F .
2. Montrer qu'elle est continue sur cet ensemble.
3. Montrer qu'elle est $\mathcal{C}^1$ sur $\mathbb{R}^*$.
4. Calculer F' et $F(0)$.
5. En déduire F .

57.13 Pour $x \in \mathbb{R}$, posons :

$$F(x) = \int_0^1 \frac{e^{-x^2(1+t^2)}}{1+t^2} dt \text{ et } G(x) = \left(\int_0^x e^{-t^2} dt \right)^2.$$

1. Montrer que F est de classe $\mathcal{C}^1$ sur $\mathbb{R}$ et préciser F' .
2. Montrer que G est de classe $\mathcal{C}^1$ sur $\mathbb{R}$ et préciser G' .
3. Montrer que la fonction $F+G$ est constante sur $\mathbb{R}$.
4. Déterminer $\lim_{x \rightarrow +\infty} F(x)$.
5. En déduire la valeur de l'intégrale de Gauss $I = \int_0^{+\infty} e^{-t^2} dt$.

57.14 Trouver les fonctions f dérivables sur $\mathbb{R}$ vérifiant : $\forall x \in \mathbb{R}, f'(x) + f(-x) = e^x$.

57.15 Trouver toutes les applications $f : \mathbb{R} \rightarrow \mathbb{R}$ continues sur $\mathbb{R}$ telles que $\forall (x, y) \in \mathbb{R}^2, f(x)f(y) = \int_{x-y}^{x+y} f(t)dt$.

57.16 1. Montrer, à l'aide d'une équation différentielle d'ordre 2, que pour tout $x > 0$:

$$\int_0^{+\infty} \frac{e^{-tx}}{1+t^2} dt = \int_0^{+\infty} \frac{\sin t}{t+x} dt$$

2. En déduire la valeur de $\int_0^{+\infty} \frac{\sin t}{t} dt$.

57.17

www



À l'aide des séries de Fourier déterminer les solutions de $y'' + y = |\cos(t)|$.

57.18 Le but de l'exercice est de déterminer si l'équation différentielle (E)

$$y'' + e^{it}y = 0$$

admet des solutions 2π -périodiques.

1.a. Montrer que la série trigonométrique $\sum_{n \geq 0} \frac{1}{(n!)^2} e^{int}$ converge uniformément sur $\mathbb{R}$ vers une fonction f 2π -périodique.

b. Montrer que la fonction f est de classe $\mathcal{C}^2$ et qu'elle est solution de (E).

2. Soit $g : \mathbb{R} \rightarrow \mathbb{C}$ une solution 2π -périodique de classe $\mathcal{C}^2$ de (E). Soient $\sum_{n \in \mathbb{Z}} c_n(g)e^{int}$ et $\sum_{n \in \mathbb{Z}} c_n(g'')e^{int}$ les séries de Fourier respectives de g et de g'' .

a. Exprimer $c_n(g'')$ en fonction de $c_n(g)$.

b. En utilisant que g est solution de (E), exprimer $c_n(g'')$ en fonction de $c_{n-1}(g)$.

c. En déduire que l'ensemble des solutions 2π -périodiques de (E) est l'espace vectoriel de dimension 1 engendré par la fonction f .

3. (E) possède-t-elle des solutions qui ne sont pas 2π -périodiques ?

Fonctions de plusieurs variables

Commençons par rappeler des définitions sur la topologie de $\mathbb{R}^n$, pour $n \in \mathbb{N}^*$.

Définition 58.1

Soit D un ensemble de $\mathbb{R}^n$.

- La **norme euclidienne** sur $\mathbb{R}^n$ est définie par l'application $\vec{x} = (x_1, \dots, x_n) \in \mathbb{R}^n \mapsto \|\vec{x}\| = \sqrt{\sum_{i=1}^n x_i^2} \in \mathbb{R}_+$.
- La **boule ouverte de centre** $\vec{a} \in \mathbb{R}^n$ **et de rayon** $r > 0$, notée $B(\vec{a}; r)$ est définie par $B(\vec{a}; r) = \{\vec{x} \in \mathbb{R}^n; \|\vec{x} - \vec{a}\| < r\}$.
- D est **ouvert** si pour tout $\vec{x} \in D$, il existe $r > 0$ tel que $B(\vec{x}; r) \subset D$.
- D est **fermé** si son complémentaire dans $\mathbb{R}^n$, noté $\mathbb{R}^n \setminus D$, est ouvert.
- D est **borné** s'il peut être inclus dans une boule de rayon fini.
- D est **compact** s'il est fermé et borné.
- Le point $\vec{a}$ est un point du **bord** de D s'il existe une suite $(\vec{d}^{(p)})_p$ de points de D , telle que $\lim_{p \rightarrow +\infty} \|\vec{d}^{(p)} - \vec{a}\| = 0$.
- D est **convexe** si pour tout couple $(\vec{a}; \vec{b})$ de points de D , le segment $[\vec{a}; \vec{b}]$ est inclus dans D , c'est-à-dire, pour tout $t \in [0; 1]$, $\vec{a} + t(\vec{b} - \vec{a}) \in D$.
- D est **connexe par arcs** si tout couple $(\vec{a}; \vec{b})$ de points de D peut être joint par un chemin ne sortant pas de D , c'est-à-dire, il existe une fonction $\alpha : [0; 1] \rightarrow \mathbb{R}^n$ continue sur $[0; 1]$ telle que $\alpha(0) = \vec{a}$, $\alpha(1) = \vec{b}$, et pour tout $t \in [0; 1]$, $\alpha(t) \in D$.
- D est un **voisinage** de $\vec{a} \in \mathbb{R}^n$ s'il existe $r > 0$ tel que $B(\vec{a}; r) \subset D$.

1. Fonctions numériques de plusieurs variables

Définition 58.2

Soit D un ensemble de $\mathbb{R}^n$, $n \geq 2$. Une **fonction numérique de plusieurs variables** est une fonction $f : D \rightarrow \mathbb{R}$, qui à tout vecteur $\vec{x} = (x_1, \dots, x_n) \in D$, associe un réel noté $f(\vec{x}) = f(x_1, \dots, x_n)$.

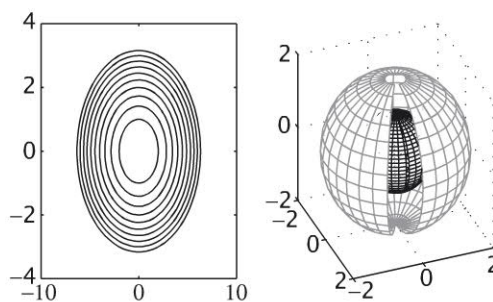
Exemple

Soit T la fonction qui à tout point de la Terre, défini par sa latitude x et sa longitude y , associe $T(x, y)$ la température moyenne en janvier en ce point. T est une fonction numérique de deux variables.

Définition 58.3

Soient D un ouvert de $\mathbb{R}^n$, $f : D \rightarrow \mathbb{R}$ une fonction numérique de plusieurs variables et $\alpha \in \mathbb{R}$. L'ensemble des points $\vec{x} \in D$ tels que $f(\vec{x}) = \alpha$ est appelé **ligne de niveau** α de la fonction f .

Si $n = 2$, nous parlons parfois de **courbes de niveau**, et si $n = 3$, nous parlons de **surfaces de niveau**.



Courbes de niveau (à gauche) de la fonction $f(x, y) = \frac{x^2}{4} + y^2$: ce sont des ellipses. Surfaces de niveau (à droite) de la fonction $f(x, y, z) = x^2 + y^2 + z^2$: ce sont des sphères.

Soient maintenant D un ouvert de $\mathbb{R}^n$ où $n \geq 1$ et $f : D \rightarrow \mathbb{R}$ une fonction numérique.

Définition 58.4

Soit $\vec{a} = (a_1, \dots, a_n)$ un point du bord de D . Nous disons que la **limite** de f quand $\vec{x} = (x_1, \dots, x_n)$ tend vers $\vec{a} = (a_1, \dots, a_n)$ est le réel L , et nous écrivons $\lim_{\vec{x} \rightarrow \vec{a}} f(\vec{x}) = L$, si les valeurs de $f(\vec{x})$ peuvent être aussi proches que nous voulons de L , en choisissant $\vec{x}$ dans un voisinage suffisamment petit de $\vec{a}$. C'est-à-dire,

$$\forall \varepsilon > 0, \exists \delta > 0, \text{ tel que si } \|\vec{x} - \vec{a}\| < \delta, \text{ alors } |f(\vec{x}) - L| < \varepsilon.$$

Exemple

Soit $f(x, y) = \frac{x^2 - y^2}{x^2 + y^2}$ une fonction de deux variables, définie sur $\mathbb{R}^2 \setminus \{(0; 0)\}$. Cette fonction n'a pas de limite lorsque (x, y) tend vers $(0; 0)$. En effet, si nous fixons $x = 0$ et faisons tendre y vers 0, nous avons $f(0, y) = -1$. La limite, si elle existe, serait donc égale à -1 . Maintenant, fixons $y = 0$ et faisons tendre x vers 0, nous avons $f(x, 0) = 1$, qui ne tend pas vers -1 !

Définition 58.5

Soit $\vec{a} = (a_1, \dots, a_n)$ un point de D . Nous disons que f est **continue** en $\vec{a}$ si $\lim_{\vec{x} \rightarrow \vec{a}} f(\vec{x}) = f(\vec{a})$. Nous disons que f est continue sur D si f est continue en tout point $\vec{a}$ de D .

Exemple

La fonction $f(x, y) = \frac{x^2 - y^2}{x^2 + y^2}$ est continue en tout point de $\mathbb{R}^2 \setminus \{(0; 0)\}$.

2. Champs vectoriels

Définition 58.6

Soit I un intervalle de $\mathbb{R}$, et soient p fonctions $f_1, \dots, f_p : I \rightarrow \mathbb{R}$ définies sur I , avec $p \geq 2$. Une **fonction vectorielle**, notée $\vec{f} : I \rightarrow \mathbb{R}^p$ est une fonction qui, à tout réel $t \in I$, associe un point de $\mathbb{R}^p$ noté $\vec{f}(t) = (f_1(t), \dots, f_p(t))$. Les fonctions $t \mapsto f_1(t), \dots, t \mapsto f_p(t)$ sont les **composantes** de $\vec{f}$. La fonction $\vec{f}$ est continue sur I si ses composantes le sont. La fonction $\vec{f}$ est de classe $\mathcal{C}^k$ sur I avec $k \in \mathbb{N}$ si ses composantes le sont.

Exemple

Soient un instant $t \in \mathbb{R}_+$ et $\vec{V}(t) = (x(t), y(t), z(t))$ la direction du vent en haut de la Tour Eiffel, à l'instant t . La fonction $\vec{V}$ est une fonction vectorielle. Nous notons aussi $\vec{V}(t) = x(t)\vec{i} + y(t)\vec{j} + z(t)\vec{k}$, où $(O, \vec{i}, \vec{j}, \vec{k})$ est le repère orthonormé de l'espace plan $\mathbb{R}^3$. Nous notons parfois $V(t) = \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix}$.

La définition suivante généralise les Définitions 58.2 et 58.6.

Définition 58.7

Soient n, p deux entiers non nuls. Soient D un ouvert de $\mathbb{R}^n$, et $f_1, \dots, f_p : D \rightarrow \mathbb{R}$, p fonctions de plusieurs variables définies sur D . La fonction $\vec{f} : D \rightarrow \mathbb{R}^p$, qui à tout vecteur $\vec{u} = (u_1, \dots, u_n) \in D$, associe le vecteur $\vec{f}(\vec{u}) = (f_1(\vec{u}), \dots, f_p(\vec{u})) \in \mathbb{R}^p$ est appelée **champ de vecteurs** ou **champ vectoriel**. Les fonctions $\vec{u} \mapsto f_1(\vec{u}), \dots, \vec{u} \mapsto f_p(\vec{u})$ sont les **composantes** de la fonction $\vec{f}$.

Définition 58.8

Soient n, p deux entiers non nuls, D un ouvert de $\mathbb{R}^n$, et $\vec{a} = (a_1, \dots, a_n)$ un point du bord de D . Un champ vectoriel $\vec{f} : D \rightarrow \mathbb{R}^p$ admet une **limite** $\vec{L} = (L_1, \dots, L_p) \in \mathbb{R}^p$ au point $\vec{a}$ si pour tout $i = 1, \dots, p$, la composante f_i admet la limite L_i au point $\vec{a}$. Nous disons que $\vec{f}$ est **continue** au point $\vec{b} = (b_1, \dots, b_n) \in D$, si ses composantes le sont. Nous disons que $\vec{f}$ est continue sur D si ses composantes le sont.

Application

Une fonction de plusieurs variables

Soit la fonction $f : (x, y) \mapsto f(x, y) = \frac{x^3 + y^3}{x^2 + y^2}$. Quel est l'ensemble de $\mathbb{R}^2$ sur lequel f est définie ? La fonction f est-elle continue sur cet ensemble ?

La fonction f est définie sur $D = \mathbb{R}^2 \setminus \{(0; 0)\}$. En tout point de D , f est continue, car

$$\forall (x_0, y_0) \in D, \lim_{(x, y) \rightarrow (x_0, y_0)} f(x, y) = \lim_{(x, y) \rightarrow (x_0, y_0)} \frac{x^3 + y^3}{x^2 + y^2} = \frac{x_0^3 + y_0^3}{x_0^2 + y_0^2} = f(x_0, y_0).$$

Regardons ce qu'il se passe si $(x, y) \rightarrow (0; 0)$.

Soit le repère orthonormé $(O, \vec{i}, \vec{j})$ du plan. Pour tout couple (x, y) de $\mathbb{R}^2$, nous associons le point M du plan, de coordonnées $\begin{pmatrix} x \\ y \end{pmatrix}$.

Observons tout d'abord le comportement de $f(x, y)$ si nous approchons (x, y) de $(0; 0)$:

- posons $x = y = t$, nous avons $f(x, y) = f(t, t) = 2t \rightarrow 0$ si $t \rightarrow 0$,
- posons $x = 0$, nous avons $f(0, y) = y \rightarrow 0$ si $y \rightarrow 0$,
- posons $y = 0$, nous avons $f(x, 0) = x \rightarrow 0$ si $x \rightarrow 0$.



Un seul de ces trois items suffit à affirmer que, si la limite de f en $(0; 0)$ existe, cette limite est nécessairement nulle. Mais ceci ne suffit pas à montrer l'existence de la limite de f en $(0; 0)$: il faut que, quel que soit le chemin pris par (x, y) pour arriver en $(0; 0)$, la limite de f soit nulle. Nous n'avons considéré ici que trois chemins conduisant au point $(0; 0)$: la première bissectrice, l'axe des ordonnées, et l'axe des abscisses.

Pour étudier le comportement de la fonction f au voisinage de $(0; 0)$, posons le changement de variables suivant : $x = r \cos \theta$, $y = r \sin \theta$ où $r = \sqrt{x^2 + y^2} \geq 0$ et $\theta = (\vec{i}; \vec{OM}) \in [0; 2\pi[$.

Si nous remplaçons l'expression de (x, y) dans f , nous obtenons :

$$f(x, y) = f(r \cos \theta, r \sin \theta) = \frac{r^3 \cos^3 \theta + r^3 \sin^3 \theta}{r^2} = r(\cos^3 \theta + \sin^3 \theta).$$

Nous avons alors cette inégalité :

$$\forall (x, y) \in D, \text{ en posant } r = \sqrt{x^2 + y^2}, \text{ nous avons } |f(x, y)| \leq 2r = 2\sqrt{x^2 + y^2},$$

c'est-à-dire

$$-2\sqrt{x^2 + y^2} \leq f(x, y) \leq 2\sqrt{x^2 + y^2}.$$

Faisons alors tendre (x, y) vers $(0; 0)$ dans cette inégalité. Nous obtenons :

$$0 = \lim_{(x,y) \rightarrow (0;0)} -2\sqrt{x^2 + y^2} \leq \lim_{(x,y) \rightarrow (0;0)} f(x, y) \leq \lim_{(x,y) \rightarrow (0;0)} 2\sqrt{x^2 + y^2} = 0.$$

Cet encadrement nous dit que la limite de f en $(0; 0)$ **existe** et vaut 0.

Ainsi, nous pouvons **prolonger f par continuité** sur $\mathbb{R}^2$ tout entier. La fonction f , ainsi prolongée en $(0; 0)$ s'écrit :

$$f : \begin{cases} \mathbb{R}^2 & \rightarrow \mathbb{R} \\ (x, y) & \mapsto \frac{x^3 + y^3}{x^2 + y^2} \text{ si } (x, y) \neq (0; 0) \\ (0; 0) & \mapsto 0 \end{cases}.$$

C'est maintenant une fonction continue sur $\mathbb{R}^2$ tout entier.

Pour s'entraîner (solutions p. 534)

58.1 Montrer que les courbes de niveau de la fonction $f(x, y) = \frac{x^2}{4} + y^2$ sont des ellipses.

58.2 Montrer que les surfaces de niveau de la fonction $f(x, y, z) = x^2 + y^2 + z^2$ sont des sphères.

58.3 Soit f une fonction continue sur le disque unité. Supposons que $\lim_{(x,y) \rightarrow (0;0)} f(x, y) = 7$. Que dire de la valeur de $f(0; 0)$?

58.4 Nous étudions la température $T(x, y, t)$ comme une fonction de la latitude

x , la longitude y et l'instant t . Nous supposons que T est définie sur $D = [-90; 90] \times [-180; 180] \times [t_0; t_1]$ où $[t_0; t_1]$ désigne la fenêtre temporelle sur laquelle nous étions la température. La fonction T est-elle, à votre avis, continue sur D ?

58.5 Calculer la limite, si elle existe, ou démontrer qu'elle n'existe pas.

1. $\lim_{(x,y) \rightarrow (2;0)} f(x, y)$ où $f(x, y) = 4x^2 - e^y$.

2. $\lim_{(x,y) \rightarrow (1;1)} f(x, y)$ où $f(x, y) = \frac{1}{\sin(2\pi xy)}$.

Dérivées partielles

Gradient – Différentielle

Nous nous intéressons dans cette Fiche aux fonctions numériques de plusieurs variables.

1. Différentiabilité des fonctions numériques de plusieurs variables

Définition 59.1

Soient D un ouvert de $\mathbb{R}^n$, avec $n \geq 2$, et $f : D \rightarrow \mathbb{R}$ une fonction numérique de plusieurs variables. Soit $\vec{a} = (a_1, \dots, a_n)$ un point de D .

1. Soit k un entier entre 1 et n , la **dérivée partielle de f par rapport à la $k^{\text{ème}}$ variable**, notée $\frac{\partial f}{\partial x_k}(\vec{a})$ est la limite, si elle existe, du taux d'accroissement :

$$\frac{\partial f}{\partial x_k}(\vec{a}) = \lim_{h \rightarrow 0} \frac{f(a_1, \dots, a_{k-1}, a_k + h, a_{k+1}, \dots, a_n) - f(a_1, \dots, a_{k-1}, a_k, a_{k+1}, \dots, a_n)}{h}.$$

La $k^{\text{ème}}$ dérivée partielle de f est donc la dérivée de la fonction $t \mapsto f(x_1, \dots, x_{k-1}, t, x_{k+1}, \dots, x_n)$ par rapport à la variable t , les autres variables étant provisoirement fixées.

2. Si toutes les dérivées partielles de la fonction f existent au point $\vec{a} \in D$ (resp. en tout point $\vec{a} \in D$), nous disons que f est **différentiable** au point $\vec{a}$ (resp. sur D).
3. Si la fonction f est différentiable sur D , nous définissons les **fonctions dérivées partielles** $\frac{\partial f}{\partial x_k} : \begin{cases} D \rightarrow \mathbb{R} \\ \vec{a} \mapsto \frac{\partial f}{\partial x_k}(\vec{a}) \end{cases}$ pour $k = 1, \dots, n$, et la **fonction gradient** :

$$\vec{\nabla} f : \begin{cases} D \rightarrow \mathbb{R}^n \\ \vec{a} \mapsto \vec{\nabla} f(\vec{a}) = \left(\frac{\partial f}{\partial x_1}(\vec{a}), \dots, \frac{\partial f}{\partial x_n}(\vec{a}) \right) \end{cases}$$

La fonction gradient est un champ vectoriel, appelé également **champ de gradient**.

4. La **fonction f est de classe $\mathcal{C}^1$ au point $\vec{a} \in D$ (resp. de classe $\mathcal{C}^1$ sur D)** si les fonctions dérivées partielles $\frac{\partial f}{\partial x_k}$, pour $k = 1, \dots, n$ existent et sont continues au point $\vec{a}$ de D (resp. sur D).

Exemple

Considérons la fonction de deux variables définie par $f(x, y) = \frac{e^{xy}}{y}$. L'ensemble de définition de cette fonction est $D = \mathbb{R} \times \mathbb{R}^*$. La fonction f admet une dérivée partielle par rapport à x et y en tout point de D : $\forall (x, y) \in D$, $\frac{\partial f}{\partial x}(x, y) = e^x$, $\frac{\partial f}{\partial y}(x, y) = \frac{xe^{xy}}{y} - \frac{e^{xy}}{y^2}$. Par définition, pour tout $(x, y) \in D$, le gradient de f s'écrit $\vec{\nabla} f(x, y) = \left(e^x, \frac{xe^{xy}}{y} - \frac{e^{xy}}{y^2} \right)$.

Définition 59.2

Soit $f : D \rightarrow \mathbb{R}$ une fonction différentiable sur D un ouvert de $\mathbb{R}^n$. La **fonction différentielle** ou **différentielle** de f en tout point $\vec{a} = (a_1, \dots, a_n)$ de D est définie par

$$df(\vec{a}) = \frac{\partial f}{\partial x_1}(\vec{a})dx_1 + \frac{\partial f}{\partial x_2}(\vec{a})dx_2 + \dots + \frac{\partial f}{\partial x_n}(\vec{a})dx_n = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\vec{a})dx_i.$$

La différentielle de f en $\vec{a}$ est parfois notée $Df(\vec{a})$.



- Cette notion de différentielle d'une fonction différentiable au point $\vec{a}$ de D signifie que si nous perturbons le point $\vec{a}$, en le remplaçant par $\vec{a} + \vec{dx} = (a_1 + dx_1, \dots, a_n + dx_n)$, alors nous avons l'approximation linéaire suivante :

$$f(\vec{a} + \vec{dx}) \approx f(\vec{a}) + \left(\frac{\partial f}{\partial x_1}(\vec{a})dx_1 + \frac{\partial f}{\partial x_2}(\vec{a})dx_2 + \dots + \frac{\partial f}{\partial x_n}(\vec{a})dx_n \right).$$

Ce que nous pouvons aussi écrire,

$$f(\vec{a} + \vec{dx}) \approx f(\vec{a}) + \vec{\nabla} f(\vec{a}) \cdot \vec{dx},$$

où $\cdot$ désigne le produit scalaire. Cette formule est parfois appelée **formule de Taylor d'ordre 1**.

- Si $n = 1$, différentiable veut dire dérivable. Si la fonction $t \mapsto f(t)$ est dérivable en t , alors sa différentielle est $f'(t)dt$ et nous avons l'approximation linéaire $f(t + dt) \approx f(t) + f'(t)dt$.

Exemple

La différentielle de la fonction définie sur $D = \mathbb{R} \times \mathbb{R}^*$ par $f(x, y) = \frac{e^{xy}}{y}$ est $df(x, y) = e^x dx + \left(\frac{x e^{xy}}{y} - \frac{e^{xy}}{y^2} \right) dy$. Au point $\vec{a} = (0; 1)$, elle vaut $df(\vec{a}) = dx - dy$.

2. Opérations sur la différentiabilité**Théorème 59.1**

Soient $f, g : D \rightarrow \mathbb{R}$ deux fonctions différentiables sur D un ouvert de $\mathbb{R}^n$. Les fonctions $f + g$ et $f \times g$ sont alors différentiables sur D et leurs différentielles, en tout point $\vec{a} = (a_1, \dots, a_n)$ de D sont définies par :

$$d(f + g)(\vec{a}) = df(\vec{a}) + dg(\vec{a}),$$

$$d(f \times g)(\vec{a}) = g(\vec{a}) \times df(\vec{a}) + f(\vec{a}) \times dg(\vec{a}).$$

Donnons maintenant la **règle de dérivation des fonctions composées**, d'abord dans deux cas simples, puis dans le cas général.

Théorème 59.2

Soit $f : \begin{cases} D \rightarrow \mathbb{R} \\ (x_1, x_2) \mapsto f(x_1, x_2) \end{cases}$ une fonction différentiable sur un ouvert D de $\mathbb{R}^2$. Remplaçons les variables x_1, x_2 par $x_1 = x_1(t)$ et $x_2 = x_2(t)$, deux fonctions dérivables définies sur un intervalle I de $\mathbb{R}$. Alors la fonction $g : \begin{cases} I \rightarrow \mathbb{R} \\ t \mapsto f(x_1(t), x_2(t)) \end{cases}$ est dérivable

sur I et sa dérivée est

$$g'(t) = \frac{\partial f}{\partial x_1} x'_1(t) + \frac{\partial f}{\partial x_2} x'_2(t).$$

Théorème 59.3

Soit $f : \begin{cases} D \longrightarrow \mathbb{R} \\ (x_1, x_2) \longmapsto f(x_1, x_2) \end{cases}$ une fonction différentiable sur un ouvert D de $\mathbb{R}^2$. Remplaçons les variables x_1, x_2 par $x_1(u, v)$, et $x_2(u, v)$ deux fonctions différentiables sur un ouvert $E \subset \mathbb{R}^2$. Alors la fonction $g : \begin{cases} E \longrightarrow \mathbb{R} \\ (u, v) \longmapsto f(x_1(u, v), x_2(u, v)) \end{cases}$ est différentiable sur E et sa différentielle est

$$dg(u, v) = \left(\frac{\partial f}{\partial x_1} \frac{\partial x_1}{\partial u} + \frac{\partial f}{\partial x_2} \frac{\partial x_2}{\partial u} \right) du + \left(\frac{\partial f}{\partial x_1} \frac{\partial x_1}{\partial v} + \frac{\partial f}{\partial x_2} \frac{\partial x_2}{\partial v} \right) dv.$$

Exemple

Soit f une fonction $(x, y) \mapsto f(x, y)$ de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$. Par le changement de variables $x = 3u + 4v$, $y = u - 3v$, nous obtenons la fonction $g : \mathbb{R}^2 \longrightarrow \mathbb{R}$ définie par : $g(u, v) = f(3u + 4v, u - 3v)$.

La fonction g est de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$ et ses dérivées partielles sont $\frac{\partial g}{\partial u}(u, v) = 3 \frac{\partial f}{\partial x}(3u + 4v, u - 3v) + \frac{\partial f}{\partial y}(3u + 4v, u - 3v)$, et $\frac{\partial g}{\partial v}(u, v) = 4 \frac{\partial f}{\partial x}(3u + 4v, u - 3v) - 3 \frac{\partial f}{\partial y}(3u + 4v, u - 3v)$.

Théorème 59.4 Règle générale de dérivation des fonctions composées

Soit $f : \begin{cases} D \longrightarrow \mathbb{R} \\ \vec{x} = (x_1, \dots, x_n) \longmapsto f(\vec{x}) \end{cases}$ une fonction différentiable sur D un ouvert de $\mathbb{R}^n$. Supposons que chaque variable x_i , pour $i = 1, \dots, n$ est remplacée par une fonction différentiable sur un ouvert $E \subset \mathbb{R}^m$, notée $x_i(t_1, \dots, t_m)$. Autrement dit, nous avons la fonction de changement de variables :

$$\begin{cases} E \subset \mathbb{R}^m \longrightarrow D \subset \mathbb{R}^n \\ \vec{t} = (t_1, \dots, t_m) \longmapsto \vec{x}(\vec{t}) = (x_1(\vec{t}), \dots, x_n(\vec{t})) \end{cases}.$$

Alors la fonction $g : \begin{cases} E \longrightarrow \mathbb{R} \\ \vec{t} = (t_1, \dots, t_m) \longmapsto f(x_1(\vec{t}), \dots, x_n(\vec{t})) \end{cases}$ est différentiable sur E et sa différentielle est

$$dg(\vec{t}) = \left(\frac{\partial f}{\partial x_1} \frac{\partial x_1}{\partial t_1} + \dots + \frac{\partial f}{\partial x_n} \frac{\partial x_n}{\partial t_1} \right) dt_1 + \dots + \left(\frac{\partial f}{\partial x_1} \frac{\partial x_1}{\partial t_m} + \dots + \frac{\partial f}{\partial x_n} \frac{\partial x_n}{\partial t_m} \right) dt_m.$$

Application La dérivée directionnelle

Soit une salle, représentée par un ouvert D de $\mathbb{R}^3$, et $f : D \longrightarrow \mathbb{R}$ une fonction différentiable sur D représentant la température en tout point $\vec{x}$ de la salle D . Étudier les variations de la température si nous nous déplaçons d'un point $\vec{x}$ de la salle vers un autre point de la salle.

Si la salle possède une source de chaleur, et que nous nous en rapprochons, la température augmente. Inversement, si nous nous en éloignons, elle diminue.



La manière dont la température change dépend donc de la direction selon laquelle nous nous dirigeons à partir de $\vec{x}$. Spécifions cette direction par un vecteur $\vec{V}$.

Plus précisément, supposons que nous nous déplaçons de $\vec{x}$ vers $\vec{x} + \vec{V}$ le long de la droite joignant $\vec{x}$ à $\vec{x} + \vec{V}$. Chaque point de cette droite est alors de la forme $\vec{x} + t\vec{V}$, où $t \in \mathbb{R}$. La fonction $t \mapsto f(\vec{x} + t\vec{V})$ représente la température sur la droite joignant $\vec{x}$ à $\vec{x} + \vec{V}$. Le **taux de variation de f** sur la droite joignant $\vec{x}$ à $\vec{x} + t\vec{V}$ est défini par : $[f(\vec{x} + t\vec{V}) - f(\vec{x})]/t$. Le numérateur de ce quotient nous dit comment varie la fonction lorsque nous nous déplaçons de $\vec{x}$ à $\vec{x} + t\vec{V}$. Supposons que la limite de ce quotient existe lorsque t tend vers 0. Cette limite est appelée **dérivée directionnelle** de f au point $\vec{x}$ dans la direction $\vec{V}$, et notée $\partial_{\vec{V}} f(\vec{x})$:

$$\partial_{\vec{V}} f(\vec{x}) = \lim_{t \rightarrow 0} \frac{f(\vec{x} + t\vec{V}) - f(\vec{x})}{t}.$$

Autrement dit, $\partial_{\vec{V}} f(\vec{x})$ est la dérivée en 0, c'est-à-dire la vitesse en 0 de la fonction $t \mapsto f(\vec{x} + t\vec{V})$.

Nous pouvons exprimer la dérivée directionnelle $\partial_{\vec{V}} f(\vec{x})$ en fonction du gradient de f , au point $\vec{x}$.

Comme f est différentiable en $\vec{x}$, nous avons si t est petit : $f(\vec{x} + t\vec{V}) \approx f(\vec{x}) + \vec{\nabla} f(\vec{x}) \cdot (t\vec{V}) = f(\vec{x}) + t\vec{\nabla} f(\vec{x}) \cdot \vec{V}$ en utilisant la linéarité du produit scalaire. Ensuite,

$$\frac{f(\vec{x} + t\vec{V}) - f(\vec{x})}{t} \approx \frac{f(\vec{x}) + t\vec{\nabla} f(\vec{x}) \cdot \vec{V} - f(\vec{x})}{t} = \vec{\nabla} f(\vec{x}) \cdot \vec{V},$$

d'où $\partial_{\vec{V}} f(\vec{x}) = \vec{\nabla} f(\vec{x}) \cdot \vec{V}$.

La variation de température sera donc maximale si la direction $\vec{V}$ est dirigée selon le gradient de température $\vec{\nabla} f(\vec{x})$, c'est-à-dire si nous déplaçons perpendiculairement aux lignes de niveau de la température.

Pour s'entraîner (solutions p. 535)

59.1 Soient le champ vectoriel $\vec{r} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ défini par $\vec{r}(x, y, z) = (x, y, z)$ et la fonction numérique $r : \mathbb{R}^3 \rightarrow \mathbb{R}$ définie par $r(x, y, z) = \|\vec{r}(x, y, z)\|$.

1. Montrer que r est différentiable sur $\mathbb{R}^3 \setminus \{(0; 0; 0)\}$, et que $\vec{\nabla} r$ est un vecteur unitaire, de même direction que $\vec{r}$.

2. Soit $n \in \mathbb{N}^*$. Montrer que $\vec{\nabla}(r^n) = nr^{n-2}\vec{r}$, sur $\mathbb{R}^3 \setminus \{(0; 0; 0)\}$.

3. Montrer que la formule reste vraie pour n un entier négatif.

4. Trouver une fonction numérique f telle que $\vec{\nabla} f = \vec{r}$.

59.2 Soit la fonction $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ définie par :

$$\begin{cases} f(x, y) = \frac{xy}{\sqrt{x^2 + y^2}} & \text{si } (x, y) \neq (0; 0) \\ f(0, 0) = 0. \end{cases}$$

Montrer que cette fonction est de classe $\mathcal{C}^1$ sur $\mathbb{R}^2 \setminus \{(0; 0)\}$, et qu'elle n'est pas de classe $\mathcal{C}^1$ en $(0; 0)$.

Dérivées partielles

Gradient – Différentielle (suite)

Nous quittons maintenant le cadre des fonctions numériques de plusieurs variables de la Fiche 59, pour celui des champs vectoriels.

 Fiche 58

1. Différentiabilité des champs vectoriels

Définition 60.1

Soient n, p deux entiers non nuls, D un ouvert de $\mathbb{R}^n$, avec $n \geq 2$, et $\vec{f} = (f_1, \dots, f_p) : D \rightarrow \mathbb{R}^p$ un champ vectoriel. Soit $\vec{a} = (a_1, \dots, a_n)$ un point de D .

1. La fonction $\vec{f}$ est **différentiable** au point $\vec{a} \in D$ (resp. sur D) si ses composantes le sont.
2. Si la fonction $\vec{f}$ est différentiable au point $\vec{a}$, la **matrice jacobienne** de $\vec{f}$ au point $\vec{a}$ est la matrice des dérivées partielles définie par la matrice de taille $p \times n$:

$$J_{\vec{f}}(\vec{a}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\vec{a}) & \frac{\partial f_1}{\partial x_2}(\vec{a}) & \dots & \frac{\partial f_1}{\partial x_n}(\vec{a}) \\ \frac{\partial f_2}{\partial x_1}(\vec{a}) & \frac{\partial f_2}{\partial x_2}(\vec{a}) & \dots & \frac{\partial f_2}{\partial x_n}(\vec{a}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_p}{\partial x_1}(\vec{a}) & \frac{\partial f_p}{\partial x_2}(\vec{a}) & \dots & \frac{\partial f_p}{\partial x_n}(\vec{a}) \end{bmatrix}$$

3. Supposons que $p = n$. Le **jacobien** de $\vec{f}$ au point $\vec{a}$ est défini par le déterminant de $J_{\vec{f}}(\vec{a})$.
4. La fonction $\vec{f}$ est de **classe $\mathcal{C}^1$** au point $\vec{a} \in D$ (resp. sur D) si ses composantes le sont.
5. Supposons que $p = n$. La fonction $\vec{f}$ est un **$\mathcal{C}^1$ -difféomorphisme** sur D si elle est bijective, de classe $\mathcal{C}^1$ sur D , et si son jacobien ne s'annule jamais sur D . Dans ce cas, pour tout $\vec{a} \in D$,

$$J_{\vec{f}^{-1}}(\vec{f}(\vec{a})) = (J_{\vec{f}}(\vec{a}))^{-1}.$$

- Si nous perturbons le point $\vec{a}$, en le remplaçant par $\vec{a} + \vec{dx} = (a_1 + dx_1, \dots, a_n + dx_n)$, alors nous avons l'approximation linéaire suivante :

$$\vec{f}(\vec{a} + \vec{dx}) \approx \vec{f}(\vec{a}) + J_{\vec{f}}(\vec{a}) \times \vec{dx},$$

où $J_{\vec{f}}(\vec{a}) \times \vec{dx}$ désigne le produit matriciel entre la matrice $J_{\vec{f}}(\vec{a})$ et le vecteur $\vec{dx}$ écrit en colonne. Cette formule est parfois appelée **formule de Taylor d'ordre 1**.

- Le jacobien intervient dans le cadre des changements de variables dans les intégrales multiples.

 Fiche 65

Exemple

Soit $\vec{f} : (r, \theta) \mapsto (r \cos \theta, r \sin \theta)$ un champ défini sur $\mathbb{R}^2$. Alors $\vec{f}$ est différentiable en tout point (r, θ) de $\mathbb{R}^2$, et sa matrice jacobienne est :

$$J_{\vec{f}}(r, \theta) = \begin{bmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{bmatrix}$$

et son jacobien est égal à r .

Définition 60.2

Soient $n \in \mathbb{N}^*$, D un ouvert de $\mathbb{R}^n$, avec $n \geq 2$, et $\vec{f} = (f_1, \dots, f_n) : D \rightarrow \mathbb{R}^n$ un champ vectoriel différentiable sur D . Soit $\vec{a} = (a_1, \dots, a_n)$ un point de D . La **divergence** du champ $\vec{f}$ au point $\vec{a}$ est définie par la trace de la matrice jacobienne $J_{\vec{f}}(\vec{a})$:

$$\operatorname{div} \vec{f}(\vec{a}) = \operatorname{tr}(J_{\vec{f}}(\vec{a})) = \sum_{i=1}^n \frac{\partial f_i}{\partial x_i}(\vec{a}).$$



Pour des champs $\vec{f}$ tels que définis dans la définition précédente, la matrice jacobienne est nécessairement une matrice carrée. Il est donc possible d'en prendre la trace.

Exemple

Si $\vec{f} = (f_1, f_2, f_3) : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ est un champ vectoriel différentiable sur $\mathbb{R}^3$, alors

$$\forall (x, y, z) \in \mathbb{R}^3, \quad \operatorname{div} \vec{f}(x, y, z) = \frac{\partial f_1}{\partial x}(x, y, z) + \frac{\partial f_2}{\partial y}(x, y, z) + \frac{\partial f_3}{\partial z}(x, y, z).$$

Définition 60.3

Si $\vec{f} = (f_1, f_2, f_3)$ est un champ vectoriel de classe $\mathcal{C}^1$ sur un ouvert D de $\mathbb{R}^3$, et à valeurs dans $\mathbb{R}^3$, alors le **rotationnel** de $\vec{f}$ est le champ vectoriel défini sur D par :

$$\operatorname{rot} \vec{f} = \left(\frac{\partial f_3}{\partial y} - \frac{\partial f_2}{\partial z} \right) \vec{i} + \left(\frac{\partial f_1}{\partial z} - \frac{\partial f_3}{\partial x} \right) \vec{j} + \left(\frac{\partial f_2}{\partial x} - \frac{\partial f_1}{\partial y} \right) \vec{k}.$$

2. Inégalité des accroissements finis**Théorème 60.1 Inégalité des accroissements finis**

Soit $\vec{f} = (f_1, \dots, f_p) : D \rightarrow \mathbb{R}^p$ un champ vectoriel différentiable sur D un ouvert convexe de $\mathbb{R}^n$. S'il existe un réel $M > 0$ tel qu'en tous points $\vec{a} \in D$, $\vec{x} \in \mathbb{R}^n$, nous avons $\|J_{\vec{f}}(\vec{a}) \times \vec{x}\| \leq M \|\vec{x}\|$ où $J_{\vec{f}}(\vec{a}) \times \vec{x}$ désigne le produit matriciel entre la matrice $J_{\vec{f}}(\vec{a})$ et le vecteur $\vec{x}$ écrit en colonne, alors,

$$\forall \vec{a}, \vec{b} \in D, \quad \|\vec{f}(\vec{a}) - \vec{f}(\vec{b})\| \leq M \|\vec{a} - \vec{b}\|.$$



En particulier, si $f : D \rightarrow \mathbb{R}$ est de classe $\mathcal{C}^1$ sur D intervalle de $\mathbb{R}$, et si il existe $M > 0$ tel que $\forall x \in D, |f'(x)| \leq M$ alors

$$\forall x, y \in D \quad |f(x) - f(y)| \leq M|x - y|.$$

Voici deux corollaires du théorème de l'inégalité des accroissements finis.

Théorème 60.2

Soit $\vec{f} = (f_1, \dots, f_p) : D \longrightarrow \mathbb{R}^p$ un champ vectoriel différentiable sur D un ouvert connexe par arcs de $\mathbb{R}^n$. Si $J_{\vec{f}}(\vec{a})$ est la matrice nulle en tout point $\vec{a} \in D$ alors $\vec{f}$ est constante sur D .



En particulier, si $f : D \longrightarrow \mathbb{R}$ est de classe $\mathcal{C}^1$ sur D intervalle, et si $f' = 0$ sur D , alors f est constante sur D .

Théorème 60.3 Théorème du point fixe

Soit $\vec{f} = (f_1, \dots, f_p) : \mathbb{R}^p \longrightarrow \mathbb{R}^p$ un champ vectoriel différentiable sur $\mathbb{R}^p$. Supposons qu'il existe un réel $M > 0$ tel qu'en tous points $\vec{a}, \vec{x} \in \mathbb{R}^p$, nous avons $\|J_{\vec{f}}(\vec{a}) \times \vec{x}\| \leq M \|\vec{x}\|$. Si M vérifie $0 < M < 1$, alors la fonction $\vec{f}$ admet un unique point fixe, i.e. il existe un unique $\vec{x} \in \mathbb{R}^p$ tel que $\vec{f}(\vec{x}) = \vec{x}$.



En particulier, si $f : \mathbb{R} \longrightarrow \mathbb{R}$ est de classe $\mathcal{C}^1$ sur $\mathbb{R}$ et s'il existe $M \in]0; 1[$ tel que $\forall x \in \mathbb{R}$, $|f'(x)| \leq M$, alors f admet un unique point fixe.

Application

Une résolution d'une équation aux dérivées partielles

Soit $D =]0; +\infty[\times \mathbb{R}$. Nous cherchons toutes les fonctions g , de classe $\mathcal{C}^1$ sur D vérifiant l'équation aux dérivées partielles suivantes :

$$x \frac{\partial g}{\partial x}(x, y) + y \frac{\partial g}{\partial y}(x, y) = \frac{y}{x}.$$



Introduisons les coordonnées polaires définies, pour $(r, \theta) \in \mathbb{R}_+^* \times \left] -\frac{\pi}{2}; \frac{\pi}{2} \right[$, par

$$x = r \cos \theta \text{ et } y = r \sin \theta,$$

et la fonction f définie par $f(r, \theta) = g(r \cos \theta, r \sin \theta)$.

Autrement dit, nous avons $g(x, y) = f(r, \theta)$.

La matrice jacobienne du changement de variables $(r, \theta) \longmapsto (x, y) = (r \cos \theta, r \sin \theta)$ est égale à :

$$J(r, \theta) = \begin{bmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{bmatrix}.$$

Le changement de variables est un $\mathcal{C}^1$ -difféomorphisme. La matrice jacobienne du changement de variables inverse $(x, y) \longmapsto (r, \theta)$ est alors égale à :

$$J^{-1}(r, \theta) = \frac{1}{r} \begin{bmatrix} r \cos \theta & r \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix}$$

Autrement dit, nous avons

$$\frac{1}{r} \begin{bmatrix} r \cos \theta & r \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} = \begin{bmatrix} \frac{\partial r}{\partial x} & \frac{\partial r}{\partial y} \\ \frac{\partial \theta}{\partial x} & \frac{\partial \theta}{\partial y} \end{bmatrix}.$$

Nous pouvons alors appliquer le Théorème 59.3, avec le changement de variables $(x, y) \longmapsto (r, \theta)$ et nous avons :

$$\begin{aligned} \frac{\partial g}{\partial x} &= \frac{\partial f}{\partial r} \frac{\partial r}{\partial x} + \frac{\partial f}{\partial \theta} \frac{\partial \theta}{\partial x} = \cos \theta \frac{\partial f}{\partial r} - \frac{\sin \theta}{r} \frac{\partial f}{\partial \theta} \\ \frac{\partial g}{\partial y} &= \frac{\partial f}{\partial r} \frac{\partial r}{\partial y} + \frac{\partial f}{\partial \theta} \frac{\partial \theta}{\partial y} = \sin \theta \frac{\partial f}{\partial r} + \frac{\cos \theta}{r} \frac{\partial f}{\partial \theta} \end{aligned}$$

La fonction g vérifie l'équation aux dérivées partielles si, et seulement si, la fonction f vérifie :

$$r \cos \theta \left(\cos \theta \frac{\partial f}{\partial r} - \frac{\sin \theta}{r} \frac{\partial f}{\partial \theta} \right) + r \sin \theta \left(\sin \theta \frac{\partial f}{\partial r} + \frac{\cos \theta}{r} \frac{\partial f}{\partial \theta} \right) = \tan \theta,$$

c'est-à-dire

$$\frac{\partial f}{\partial r} = \frac{\tan \theta}{r}.$$

Il suffit alors d'intégrer cette dernière équation par rapport à r . Nous obtenons $f(r, \theta) = \ln r \times \tan \theta + \kappa(\theta)$ où κ est une fonction quelconque de classe $\mathcal{C}^1$ sur $\left] -\frac{\pi}{2}; \frac{\pi}{2} \right[$. Finalement, les solutions cherchées sont

$$g(x, y) = \frac{y}{2x} \ln(x^2 + y^2) + \kappa\left(\arctan \frac{y}{x}\right).$$

Pour s'entraîner (solutions p. 535)

60.1 Montrer que la fonction $\vec{f} : \mathbb{R}_+^* \times]0; 2\pi[\rightarrow \mathbb{R}^2$, telle que $\vec{f}(r, \theta) = (r \cos \theta, r \sin \theta)$ est un $\mathcal{C}^1$ -difféomorphisme sur $\mathbb{R}_+^* \times]0; 2\pi[$.

60.2 Soit la fonction $\vec{f} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, définie par $f(x, y, z) = \left(\frac{x}{2}, \frac{y}{3}, \frac{z}{4}\right)$. Montrer que $\vec{f}$ est de classe $\mathcal{C}^1$ sur $\mathbb{R}^3$, et qu'elle a un unique point fixe, à déterminer.

60.3 Soit $\vec{f} = (f_1, f_2) : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ un champ vectoriel à valeurs dans $\mathbb{R}^2$, où f_1 et f_2 sont deux fonctions numériques définies sur $\mathbb{R}^2$.

Nous supposons que $\vec{f}$ dérive d'un potentiel φ sur $\mathbb{R}^2$, c'est-à-dire $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}$ est une fonction numérique différentiable sur $\mathbb{R}^2$ telle que $\vec{f} = \vec{\nabla} \varphi$ sur $\mathbb{R}^2$.

Soit $u : D \subset \mathbb{R} \rightarrow \mathbb{R}$ une fonction solution de l'équation différentielle suivante :

$$f_2(t, u(t)) \times u'(t) + f_1(t, u(t)) = 0, \quad t \in \mathbb{R}.$$

1. Montrer que si u est solution de l'équation différentielle, alors u vérifie :

$$\varphi(t, u(t)) = \text{constante}, \quad \forall t \in \mathbb{R}.$$

2. En déduire une solution de :

$$\begin{cases} u^2(t) \times u'(t) + \sin t = 0 \\ u(0) = 3 \end{cases}$$

Indications pour la question 2 : explicitez la fonction $\vec{f}$, vérifiez qu'elle dérive d'un potentiel φ .

60.4 **1.** Soit le champ vectoriel suivant $\vec{f}(x_1, x_2) = \left(\frac{1}{2} \sin(x_1 + x_2), \frac{1}{2} \cos(x_1 - x_2)\right)$. Montrer qu'il est de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$.

2. Soient $\vec{a} = (a_1, a_2)$, $\vec{x} = (x_1, x_2) \in \mathbb{R}^2$. Montrer que $J_{\vec{f}}(\vec{a}) \times \vec{x} = \frac{1}{2} \begin{pmatrix} (x_1 + x_2) \cos(a_1 + a_2) \\ -(x_1 - x_2) \sin(a_1 - a_2) \end{pmatrix}$

3. En déduire que pour tout $\vec{a}, \vec{x} \in \mathbb{R}^2$, nous avons $\|J_{\vec{f}}(\vec{a}) \times \vec{x}\|^2 \leq \frac{1}{2} \|\vec{x}\|^2$.

4. Montrer que pour tout $\vec{a}, \vec{b} \in \mathbb{R}^2$, nous avons $\|\vec{f}(\vec{a}) - \vec{f}(\vec{b})\| \leq \frac{1}{\sqrt{2}} \|\vec{a} - \vec{b}\|$.

5. En déduire que le système d'équations $x_1 = \frac{1}{2} \sin(x_1 + x_2)$, $x_2 = \frac{1}{2} \cos(x_1 - x_2)$ admet au plus une solution.

Dérivées partielles d'ordre deux

Optimisation

Nous avons étudié à la Fiche 37 la notion de dérivée seconde d'une fonction numérique définie sur un intervalle de $\mathbb{R}$. Nous étendons cette notion aux cas des fonctions numériques de plusieurs variables.

Définition 61.1

Soient D un ouvert de $\mathbb{R}^n$ et $f : D \rightarrow \mathbb{R}$ une fonction possédant des dérivées partielles en tout point de D .

- Si les fonctions dérivées partielles $\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n}$ admettent en $\vec{a} \in D$ des dérivées partielles d'ordre un, nous pouvons alors définir des **dérivées partielles d'ordre deux** ou **dérivées partielles secondes** de f en $\vec{a}$:

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(\vec{a}) = \frac{\partial}{\partial x_i} \left(\frac{\partial f}{\partial x_j} \right)(\vec{a}), \quad \text{pour } i, j = 1, \dots, n.$$

Dans le cas où $i = j$, nous notons plutôt $\frac{\partial^2 f}{\partial x_i^2}(\vec{a}) = \frac{\partial}{\partial x_i} \left(\frac{\partial f}{\partial x_i} \right)(\vec{a})$.

- Si toutes les dérivées partielles secondes de la fonction f existent au point $\vec{a} \in D$ (resp. en tout point $\vec{a} \in D$), nous disons que f est **deux fois différentiable** au point $\vec{a}$ (resp. sur D).
- Si la fonction f est deux fois différentiable sur D , nous définissons les **fonctions dérivées partielles secondes** $\frac{\partial^2 f}{\partial x_i \partial x_j} : \begin{cases} D \rightarrow \mathbb{R} \\ \vec{a} \mapsto \frac{\partial^2 f}{\partial x_i \partial x_j}(\vec{a}) \end{cases}$ pour $i, j = 1, \dots, n$, et la **matrice hessienne** de f en tout point $\vec{a} \in D$:

$$H_f(\vec{a}) = \left[\frac{\partial^2 f}{\partial x_i \partial x_j}(\vec{a}) \right]_{1 \leq i, j \leq n} = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2}(\vec{a}) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(\vec{a}) & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(\vec{a}) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(\vec{a}) & \frac{\partial^2 f}{\partial x_2^2}(\vec{a}) & \dots & \frac{\partial^2 f}{\partial x_2 \partial x_n}(\vec{a}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(\vec{a}) & \frac{\partial^2 f}{\partial x_n \partial x_2}(\vec{a}) & \dots & \frac{\partial^2 f}{\partial x_n^2}(\vec{a}) \end{bmatrix}$$

- La fonction f est de classe $\mathcal{C}^2$ au point $\vec{a} \in D$ (resp. de classe $\mathcal{C}^2$ sur D) si les fonctions dérivées partielles secondes $\frac{\partial^2 f}{\partial x_i \partial x_j}$, pour $i, j = 1, \dots, n$ existent et sont continues au point $\vec{a}$ de D (resp. sur D).

Théorème 61.1 Théorème de Schwarz

Soient D un ouvert de $\mathbb{R}^n$ et $f : D \rightarrow \mathbb{R}$ une fonction de classe $\mathcal{C}^2$ sur D . Alors, pour tout $i, j = 1, \dots, n$, pour tout $\vec{a} \in D$, nous avons $\frac{\partial^2 f}{\partial x_i \partial x_j}(\vec{a}) = \frac{\partial^2 f}{\partial x_j \partial x_i}(\vec{a})$, c'est-à-dire la matrice hessienne de f est une matrice symétrique.

Exemple

La fonction définie par $f(x, y) = x^3 \cos(2xy) + y^2$ est de classe $\mathcal{C}^2$ sur $\mathbb{R}^2$. Nous avons $\frac{\partial^2 f}{\partial y \partial x} = -8x^3 \sin(2xy) - 4x^4 y \cos(2xy) = \frac{\partial^2 f}{\partial x \partial y}$.

Définition 61.2

Soit une fonction $f : D \rightarrow \mathbb{R}$ définie sur D un ouvert de $\mathbb{R}^n$.

- La fonction f admet un **maximum local** en $\vec{a} \in D$ si $f(\vec{x}) \leq f(\vec{a})$ pour tout point $\vec{x}$ dans un voisinage de $\vec{a}$.
- La fonction f admet un **minimum local** en $\vec{a} \in D$ si $f(\vec{x}) \geq f(\vec{a})$ pour tout point $\vec{x}$ dans un voisinage de $\vec{a}$.
- La fonction f admet un **maximum global** en $\vec{a} \in D$ si $f(\vec{x}) \leq f(\vec{a})$ pour tout point $\vec{x} \in D$.
- La fonction f admet un **minimum global** en $\vec{a} \in D$ si $f(\vec{x}) \geq f(\vec{a})$ pour tout point $\vec{x} \in D$.
- Un point $\vec{a} \in D$ est un **point stationnaire** de f si f est différentiable en $\vec{a}$ et si $\vec{\nabla} f(\vec{a}) = \vec{0}$.
- Un point $\vec{a} \in D$ est un **point selle** de f si c'est un point stationnaire, et si dans tout voisinage de $\vec{a}$, il existe des points $\vec{x}$ tels que $f(\vec{x}) > f(\vec{a})$ et des points $\vec{x}$ tels que $f(\vec{x}) < f(\vec{a})$.
- Un **extremum local** (resp. **global**) de la fonction f est un maximum ou un minimum local (resp. global) de la fonction f .

Nous donnons ci-dessous une condition nécessaire d'existence d'un extremum local.

Théorème 61.2

Soit $f : D \rightarrow \mathbb{R}$ une fonction différentiable sur un ouvert D de $\mathbb{R}^n$. Si f admet un extremum local au point $\vec{a}$, alors la différentielle de f est nulle au point $\vec{a}$.

Le théorème ci-dessous permet de reconnaître si un point stationnaire est un extremum local ou un point selle, dans le cas des fonctions de deux variables.

Théorème 61.3

Soient $f : \begin{cases} D \rightarrow \mathbb{R} \\ (x, y) \mapsto f(x, y) \end{cases}$ une fonction de classe $\mathcal{C}^2$ sur un ouvert D de $\mathbb{R}^2$, et $\vec{a} \in D$

un point stationnaire de f . Nous posons : $r = \frac{\partial^2 f}{\partial x^2}(\vec{a})$, $t = \frac{\partial^2 f}{\partial y^2}(\vec{a})$, $s = \frac{\partial^2 f}{\partial x \partial y}(\vec{a})$.

- Si $s^2 - rt < 0$ et $r > 0$ alors f admet un minimum local en $\vec{a}$.
- Si $s^2 - rt < 0$ et $r < 0$ alors f admet un maximum local en $\vec{a}$.
- Si $s^2 - rt > 0$ alors $\vec{a}$ n'est pas un extremum local, mais un point selle.

Le théorème suivant stipule que les fonctions continues sur un ensemble compact admettent un maximum global et un minimum global.

Théorème 61.4

Soit $f : C \rightarrow \mathbb{R}$ une fonction différentiable sur un ensemble compact C de $\mathbb{R}^n$. Alors f admet un maximum global et un minimum global : il existe $\vec{a}, \vec{b} \in C$ tels que

$$f(\vec{a}) = \min_{\vec{x} \in C} f(\vec{x}), \quad \text{et} \quad f(\vec{b}) = \max_{\vec{x} \in C} f(\vec{x}).$$

Application Une recherche d'extremum – La piscine

1. Soit f la fonction de $\mathbb{R}^2$ dans $\mathbb{R}$ définie par : $\forall (x, y) \in \mathbb{R}^2, f(x, y) = x^2 + xy + y^2 - 3y + 2$. Montrer que f admet un extremum local.

2. Un fournisseur nous donne $S = 12 \text{ m}^2$ de carrelage pour carrelé l'intérieur d'une piscine rectangulaire. Quel est le volume maximal de la piscine ?

1. La fonction f est de classe $\mathcal{C}^2$ sur $\mathbb{R}^2$ car c'est une fonction polynômiale en x, y . Nous avons : $\frac{\partial f}{\partial x} = 2x + y$ et $\frac{\partial f}{\partial y} = x + 2y - 3$. Le point $M(x, y)$ est un extremum de f si $\vec{\nabla} f(x, y) = \vec{0}$, c'est-à-dire

$$\begin{cases} 2x + y = 0 \\ x + 2y - 3 = 0 \end{cases}$$

Le déterminant du système est non nul, car

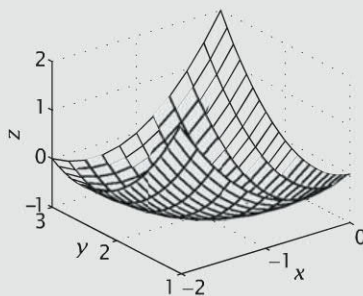
$$\begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix} = 3,$$

la solution de ce système est donc unique.

Cette solution est $x = \frac{\begin{vmatrix} 0 & 1 \\ 3 & 2 \end{vmatrix}}{\begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix}} = -1$ et $y = \frac{\begin{vmatrix} 2 & 0 \\ 1 & 3 \end{vmatrix}}{\begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix}} = 2$. La fonction f admet donc un unique point stationnaire, le point $A(-1; 2)$.

Cherchons maintenant si ce point stationnaire est un extremum de f .

Nous avons : $r = \frac{\partial^2 f}{\partial x^2}(-1; 2) = 2$, $t = \frac{\partial^2 f}{\partial y^2}(-1; 2) = 2$, et $s = \frac{\partial^2 f}{\partial y \partial x}(-1; 2) = 1$. Ainsi, $s^2 - rt = -3 < 0$, donc A est un extremum local, et puisque $r > 0$, A est un minimum local.



Le point $A(-1; 2)$ est un minimum local de la fonction $f(x, y) = x^2 + xy + y^2 - 3y + 2$.

2. Notons x, y, z la longueur, la largeur et la profondeur de la piscine. Soit S la surface à carrelé de la piscine. Nous avons $xy + 2xz + 2yz = S$, ainsi, z peut s'exprimer en fonction de x et y .



Ensuite, nous cherchons à maximiser une fonction $V(x, y)$ donnant le volume de la piscine, et ne dépendant que des variables x et y .

Nous avons $z = \frac{S - xy}{2(x + y)}$, et le volume à maximiser est $V(x, y) = xy \frac{S - xy}{2(x + y)}$. La fonction V est de classe $\mathcal{C}^1$ sur $(\mathbb{R}_+^*)^2$. Nous cherchons (x, y) tel que $\frac{\partial V}{\partial x} = \frac{\partial V}{\partial y} = 0$, c'est-à-dire

$$\begin{cases} \frac{\partial V}{\partial x} = \frac{Sy - 2xy^2}{2(x + y)} - \frac{Sxy - x^2y^2}{2(x + y)^2} = 0 \\ \frac{\partial V}{\partial y} = \frac{Sx - 2x^2y}{2(x + y)} - \frac{Sxy - x^2y^2}{2(x + y)^2} = 0. \end{cases}$$

Faisons $\frac{\partial V}{\partial x} - \frac{\partial V}{\partial y} = 0$, nous obtenons $\frac{S(y - x) - 2xy(y - x)}{2(x + y)} = 0$ ce qui implique soit $y = x$, soit $S = 2xy$.

Si nous injectons $S = 2xy$ dans $\frac{\partial V}{\partial x}$, nous n'obtenons pas $\frac{\partial V}{\partial x} = 0$ pour $(x, y) \in (\mathbb{R}_+^*)^2$. Donc, si un volume maximal existe, nous avons nécessairement $y = x$.

Maintenant, injectons $y = x$ dans $\frac{\partial V}{\partial x} = 0$, cela nous donne $\frac{Sx - 2x^3}{4x} - \frac{Sx^2 - x^4}{8x^2} = 0$, c'est-à-dire $Sx^2 - 3x^4 = 0$, donc $x = 2$, $y = 2$, et $z = 1$. Vérifions que $(2; 2)$ est un maximum local de la fonction V : nous avons $s = -\frac{1}{2}$, $r = t = -1$ donc $s^2 - rt < 0$ et $r < 0$ donc $(2; 2)$ est bien un maximum local, et le volume maximal de la piscine est $V(2; 2) = 4 \text{ m}^3$.

Pour s'entraîner (solutions p. 536)

61.1 Soit la fonction f définie sur $\mathbb{R}^2$ par : $f(x, y) = x^3 \cos(xy) + 3y^2$, calculer $\frac{\partial^2 f}{\partial x^2}$, $\frac{\partial^2 f}{\partial y^2}$, $\frac{\partial^2 f}{\partial x \partial y}$ et $\frac{\partial^2 f}{\partial y \partial x}$. Montrer que $(0; 0)$ est un point stationnaire de f . Montrer que c'est un point selle.

61.2 Soit la fonction f définie sur $\mathbb{R}^2$ par : $f(x, y) = 1 + (x + 1)^2 + (y - 1)^2$. Étudier les éventuels extrema de f .

61.3 Soit la fonction f définie sur $\mathbb{R}^2$ par : $f(x, y) = x^2 - y^2$. La fonction f admet-elle un extremum ?

61.4 Soit une corde de longueur $L = 100$ mètres. Quel est le champ rectangulaire de surface maximale que nous pouvons clôturer avec cette corde ? *Indication : notons x, y les longueurs des côtés du champ. Nous avons $2(x + y) = L$, ainsi, y peut s'exprimer en fonction de x . Ensuite, nous cherchons à maximiser une fonction $f(x)$ donnant la surface du champ, et ne dépendant que de la variable x .*

61.5 Soit la fonction f définie sur $\mathbb{R}^2$ par : $f(x, y) = x^2 - y^2$ définie sur le pavé $C = \{x \in [0; 1], y \in [0; 1]\}$. Pourquoi la fonction f admet-elle un maximum et un minimum sur C ? Les déterminer.

Courbes et surfaces paramétrées

Les courbes et surfaces paramétrées jouent un rôle important dans toutes les sciences physiques. Nous les présentons dans cette Fiche. Nous notons $(O, \vec{i}, \vec{j}, \vec{k})$ le repère orthonormé de $\mathbb{R}^3$.

1. Courbes paramétrées

Définition 62.1

Soient I un intervalle de $\mathbb{R}$ et $\vec{f}$ une fonction définie sur I à valeurs dans $\mathbb{R}^3$. Nous notons $\vec{f}(t) = x(t)\vec{i} + y(t)\vec{j} + z(t)\vec{k}$, pour tout $t \in I$. L'ensemble des points (x, y, z) de $\mathbb{R}^3$ tels que

$$x = x(t), \quad y = y(t), \quad z = z(t) \quad (62.1)$$

pour $t \in I$ est appelé **courbe paramétrée** (ou **arc paramétré**) Γ . Les équations (62.1) sont appelées **équations paramétriques** de Γ .



Les **courbes paramétrées planes** sont un cas particulier : $\vec{f}(t) = x(t)\vec{i} + y(t)\vec{j}$, et $z(t) = 0, \forall t \in I$.

Définition 62.2

Soient Γ une courbe paramétrée par la fonction $\vec{f}$ de classe $\mathcal{C}^1$ sur l'intervalle I , et $a \in I$.

- L'**abscisse curviligne** s est la fonction définie pour tout $t \in I$, par

$$s(t) = \int_a^t \|\vec{f}'(u)\| du = \int_a^t \sqrt{x'(u)^2 + y'(u)^2 + z'(u)^2} du.$$

- Soient $t_1, t_2 \in I$. La **longueur de la courbe entre deux points** M_1 et M_2 de la courbe Γ tels que $\vec{OM}_1 = \vec{f}(t_1)$ et $\vec{OM}_2 = \vec{f}(t_2)$ est définie par

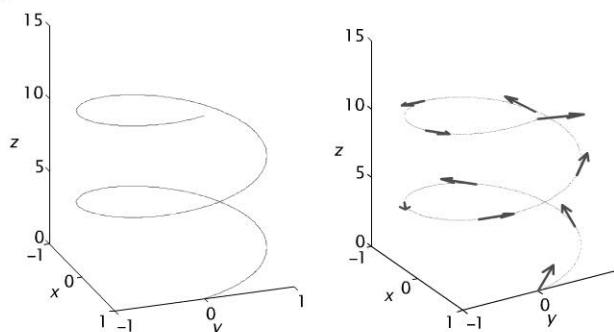
$$|s(t_2) - s(t_1)| = \left| \int_{t_1}^{t_2} \|\vec{f}'(u)\| du \right|.$$

Si $I = [a; b]$, la **longueur de la courbe** Γ est définie par $\int_a^b \|\vec{f}'(u)\| du$.

- Soient $t \in I$ et M_t défini par $\vec{OM}_t = \vec{f}(t)$ un point de la courbe Γ .
 - Le point M_t est **régulier** si $\vec{f}'(t) \neq \vec{0}$. Dans ce cas, nous définissons le **vecteur tangent unitaire** au point M_t par $\vec{T}(t) = \frac{\vec{f}'(t)}{\|\vec{f}'(t)\|}$. Une courbe est **lisse** ou **régulière** si tous ses points sont réguliers.
 - Le point M_t est **singulier** ou **stationnaire** si $\vec{f}'(t) = \vec{0}$.

Exemple

La figure suivante est le tracé de la courbe paramétrée d'équations paramétriques $x(t) = \cos t$, $y(t) = \sin t$, $z(t) = t$, pour $t \in [0; 4\pi]$. C'est une hélice. La longueur de l'hélice représentée est $4\sqrt{2}\pi$.



Exemple de courbe paramétrée. À droite, nous avons rajouté quelques vecteurs tangents à la courbe.

Définition 62.3

Soit Γ une courbe paramétrée par la fonction $\vec{f}$ de classe $\mathcal{C}^2$ sur l'intervalle I . Soit M_t un point régulier de Γ , défini par $\overrightarrow{OM_t} = \vec{f}(t)$, avec $t \in I$.

- La **courbure** de la courbe Γ au point M_t est la vitesse avec laquelle la courbe change de direction en ce point, c'est-à-dire le réel positif

$$\kappa(t) = \frac{\|\vec{T}'(t)\|}{\|\vec{f}'(t)\|} = \frac{\|\vec{f}'(t) \wedge \vec{f}''(t)\|}{\|\vec{f}'(t)\|^3}, \text{ si } \vec{f}'(t) \neq \vec{0}$$

où $\vec{T}'(t)$ est la dérivée du vecteur tangent unitaire au point M_t . Le point M_t est **birégulier** si sa courbure est non nulle.

- Le **rayon de courbure** au point M_t birégulier est défini par le réel positif

$$R(t) = \frac{1}{\kappa(t)}.$$



- La courbure et le rayon de courbure d'un cercle de rayon $r > 0$ sont constants, et valent respectivement $\frac{1}{r}$ et r . La courbure d'une droite est nulle.
- La courbure d'un cercle de faible rayon est plus grande que la courbure d'un cercle de grand rayon. De même, la courbure d'une route de montagne en forme d'épingle à cheveux est plus grande que la courbure d'une autoroute rectiligne.

► Cas particulier : les courbes planes

En particulier, si la courbe Γ est plane, paramétrée par $\vec{f}(t) = x(t)\vec{i} + y(t)\vec{j}$, pour $t \in I$, alors :

$$\|\vec{f}'(t)\| = \sqrt{x'(t)^2 + y'(t)^2} \text{ et } \kappa(t) = \frac{|x'(t)y''(t) - x''(t)y'(t)|}{(x'(t)^2 + y'(t)^2)^{\frac{3}{2}}}.$$

Si la courbe est paramétrée en polaire, par $\vec{f}(t) = r(t)\cos(t)\vec{i} + r(t)\sin(t)\vec{j}$, pour $t \in I$, alors :

$$\|\vec{f}'(t)\| = \sqrt{r(t)^2 + r'(t)^2} \text{ et } \kappa(t) = \frac{|2r'(t)^2 - r(t)r''(t) + r(t)^2|}{(r(t)^2 + r'(t)^2)^{\frac{3}{2}}}.$$

Fiche 64

2. Surfaces paramétrées

Définition 62.4

Soient D un ouvert de $\mathbb{R}^2$, et $\vec{f}$ une fonction de classe $\mathcal{C}^1$ sur D à valeurs dans $\mathbb{R}^3$. Nous notons $\vec{f}(u, v) = x(u, v)\vec{i} + y(u, v)\vec{j} + z(u, v)\vec{k}$, pour tout $(u, v) \in D$. L'ensemble des points (x, y, z) de $\mathbb{R}^3$ tels que

$$x = x(u, v), \quad y = y(u, v), \quad z = z(u, v) \quad (62.2)$$

pour $(u, v) \in D$ est appelé **surface paramétrée** (ou **nappe paramétrée**) Σ . Les équations (62.2) sont appelées **équations paramétriques** de Σ .

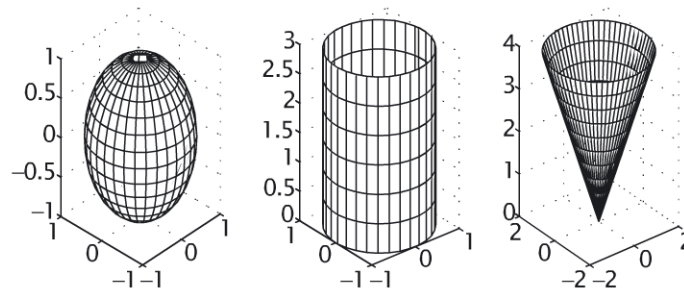
Soit M un point de la surface Σ défini par $\vec{OM} = \vec{f}(u, v)$, pour $(u, v) \in D$.

- Le point M est **régulier** si $\frac{\partial \vec{f}}{\partial u}(u, v)$ et $\frac{\partial \vec{f}}{\partial v}(u, v)$ sont des vecteurs indépendants, c'est-à-dire si $\vec{n}(u, v) = \frac{\partial \vec{f}}{\partial u}(u, v) \wedge \frac{\partial \vec{f}}{\partial v}(u, v) \neq \vec{0}$. Dans ce cas, le **plan tangent** à la surface Σ au point M est défini par le plan passant par M et dirigé par $\frac{\partial \vec{f}}{\partial u}(u, v)$ et $\frac{\partial \vec{f}}{\partial v}(u, v)$. Le vecteur $\vec{n}(u, v)$ est alors appelé **vecteur normal** à la surface au point M . Une surface est **lisse** ou **régulière** si tous ses points sont réguliers.
- Dans le cas contraire, $\frac{\partial \vec{f}}{\partial u}(u, v) \wedge \frac{\partial \vec{f}}{\partial v}(u, v) = \vec{0}$, le point M est dit **singulier**.

L'**aire** de la surface Σ se calcule par la formule :

$$A = \iint_D \left\| \frac{\partial \vec{f}}{\partial u}(u, v) \wedge \frac{\partial \vec{f}}{\partial v}(u, v) \right\| du dv.$$

 Fiche 65



Exemples de surfaces paramétrées : la sphère, le cylindre, le cône.

Exemple

L'aire de la sphère de rayon $R > 0$ paramétrée par $\vec{f}(\phi, \theta) = R \sin \phi \cos \theta \vec{i} + R \sin \phi \sin \theta \vec{j} + R \cos \phi \vec{k}$, avec $(\phi, \theta) \in]0; \pi[\times]0; 2\pi[$, est égale à : $\int_0^{2\pi} \int_0^\pi \left\| \frac{\partial \vec{f}}{\partial \phi}(\phi, \theta) \wedge \frac{\partial \vec{f}}{\partial \theta}(\phi, \theta) \right\| d\phi d\theta = 4\pi R^2$.

Application Une courbe paramétrée

Calculer la longueur et le rayon de courbure de la courbe Γ d'équations paramétriques : $x(\theta) = r(\theta) \cos \theta$ et $y(\theta) = r(\theta) \sin \theta$ avec $r(\theta) = \text{th}\left(\frac{\theta}{2}\right)$, pour $\theta \in [0; T]$ avec $T > 0$.

Fiche 40

La fonction r est de classe $\mathcal{C}^2$ sur $\mathbb{R}_+^*$, et pour tout $\theta \in \mathbb{R}_+^*$, nous avons $r'(\theta) = \frac{1}{2 \text{ch}^2\left(\frac{\theta}{2}\right)}$ et

$$r''(\theta) = \frac{-\text{sh}\left(\frac{\theta}{2}\right)}{2 \text{ch}^3\left(\frac{\theta}{2}\right)}. \text{ Ainsi,}$$

$$\|\vec{f}'(\theta)\| = \sqrt{r^2(\theta) + r'^2(\theta)} = \sqrt{\frac{\text{sh}^2\left(\frac{\theta}{2}\right)}{\text{ch}^2\left(\frac{\theta}{2}\right)} + \frac{1}{4 \text{ch}^4\left(\frac{\theta}{2}\right)}} = \frac{\text{ch} \theta}{2 \text{ch}^2\left(\frac{\theta}{2}\right)} = 1 - \frac{1}{2 \text{ch}^2\left(\frac{\theta}{2}\right)}.$$

La longueur de la courbe est donc

$$L(T) = \int_0^T \|\vec{f}'(\theta)\| d\theta = \int_0^T \left(1 - \frac{1}{2 \text{ch}^2\left(\frac{\theta}{2}\right)}\right) d\theta = \left[\theta - \text{th}\left(\frac{\theta}{2}\right)\right]_0^T = T - \text{th}\left(\frac{T}{2}\right).$$

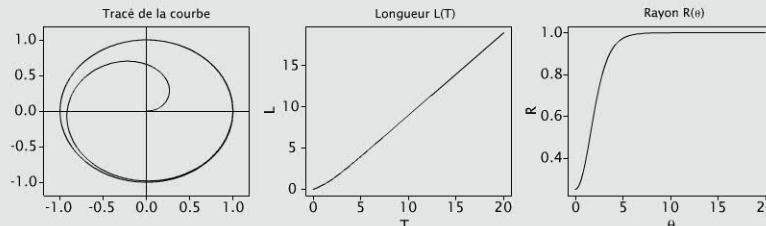
Le rayon de courbure en un point $M(\theta)$ de la courbe est égal à :

$$R(\theta) = \frac{\|\vec{f}'(\theta)\|^3}{\left|\frac{1}{2 \text{ch}^4\left(\frac{\theta}{2}\right)} + \frac{\text{th}\left(\frac{\theta}{2}\right) \text{sh}\left(\frac{\theta}{2}\right)}{2 \text{ch}^3\left(\frac{\theta}{2}\right)} + \text{th}^2\left(\frac{\theta}{2}\right)\right|} = \frac{\frac{\text{ch}^3 \theta}{2^3 \text{ch}^6\left(\frac{\theta}{2}\right)}}{\frac{1+2 \text{sh}^2\left(\frac{\theta}{2}\right)}{2 \text{ch}^2\left(\frac{\theta}{2}\right)}} = \frac{\text{ch}^2 \theta}{4 \text{ch}^4\left(\frac{\theta}{2}\right)}.$$



Le rayon de courbure tend vers 1 lorsque $\theta \rightarrow +\infty$, ce qui correspond au fait que la courbe Γ est asymptote au cercle de rayon 1.

Fiche 64



Courbe Γ : tracé, longueur et courbure en fonction de T .

Pour s'entraîner (solutions p. 536)

62.1 Calculer la longueur et le rayon de courbure d'une arche de cycloïde d'équations paramétriques :

$$x(t) = R(t - \sin(t)) \text{ et } y(t) = R(1 - \cos(t)), \text{ pour } t \in [0; 2\pi], \quad R > 0.$$

62.2 Calculer le rayon de courbure de la courbe paramétrée en polaire par $r(\theta) = a \cos^3\left(\frac{\theta}{3}\right)$, $-\frac{3\pi}{2} < \theta < \frac{3\pi}{2}$.

62.3 Calculer la longueur de la courbe paramétrée par $\vec{f}(t) = \frac{1}{\text{ch} t} \vec{i} + (t - \text{th} t) \vec{j}$, pour $t \in \mathbb{R}$.

62.4 1. Calculer la surface du cylindre (sans sa base) paramétré par $x(u, v) = R \cos u$, $y(u, v) = R \sin u$, $z(u, v) = v$, pour $(u, v) \in]0; 2\pi[\times]0; L[$, où $R, L > 0$.

2. Même question pour le cône (sans sa base) paramétré par $x(u, v) = v \cos u$, $y(u, v) = v \sin u$, $z(u, v) = 2v$, pour $(u, v) \in]0; 2\pi[\times]0; L[$, où $L > 0$.

Nous détaillons ici le cas des courbes paramétrées planes. Celles-ci sont définies par la donnée d'un intervalle I de $\mathbb{R}$ et d'une fonction $\vec{f}$ définie sur I à valeurs dans $\mathbb{R}^2$. Nous notons $\vec{f}(t) = x(t)\vec{i} + y(t)\vec{j}$, pour tout $t \in I$. La courbe est alors l'ensemble des points (x, y) du plan $(O, \vec{i}, \vec{j})$ tels que

$$x = x(t), \quad y = y(t)$$

pour $t \in I$.

1. Étude locale

Le théorème ci-dessous permet l'étude locale d'une courbe plane paramétrée.

Théorème 63.1

Soit une courbe plane paramétrée par la fonction $\vec{f}(t) = x(t)\vec{i} + y(t)\vec{j}$ définie en tout point $t \in I$ intervalle de $\mathbb{R}$. Soient $t_0 \in I$ et $M_{t_0} = \begin{pmatrix} x(t_0) \\ y(t_0) \end{pmatrix}$ un point de la courbe.

Supposons que la fonction $\vec{f}$ est suffisamment dérivable en t_0 . Nous cherchons la première dérivée de $\vec{f}$ non nulle en t_0 , et nous la notons $\vec{f}^{(p)}(t_0) = x^{(p)}(t_0)\vec{i} + y^{(p)}(t_0)\vec{j} \neq \vec{0}$, où p est un entier tel que $p \geq 1$.

L'allure de la courbe au voisinage du point M_{t_0} dépend de $\vec{f}^{(p)}(t_0)$, le premier vecteur dérivé non nul en t_0 : ce vecteur dirige la **tangente** à la courbe au point M_{t_0} . Plus précisément, nous avons les deux cas suivants, suivant que $p = 1$ (cas où M_{t_0} est un point régulier), ou $p \geq 2$ (cas où M_{t_0} est un point singulier).

1. Si M_{t_0} est un point **régulier** de la courbe, c'est-à-dire si $\vec{f}'(t_0) = x'(t_0)\vec{i} + y'(t_0)\vec{j} \neq \vec{0}$, alors la droite passant par M_{t_0} de vecteur directeur $\vec{f}'(t_0)$ est la **tangente** en M_{t_0} à la courbe.

Son équation est :

$$(x - x(t_0)) \times y'(t_0) - (y - y(t_0)) \times x'(t_0) = 0.$$

Les points M_{t_0} à tangentes horizontales sont les points où $y'(t_0) = 0$ et $x'(t_0) \neq 0$.

Les points à tangentes verticales sont les points où $x'(t_0) = 0$ et $y'(t_0) \neq 0$.

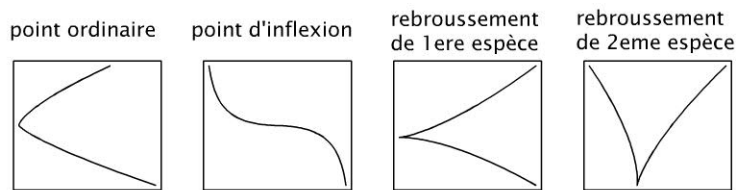
2. Si M_{t_0} est un point **singulier** de la courbe, soit $\vec{f}^{(p)}(t_0) = x^{(p)}(t_0)\vec{i} + y^{(p)}(t_0)\vec{j} \neq \vec{0}$, où $p \geq 2$, le premier vecteur dérivé non nul en t_0 . La droite passant par M_{t_0} de vecteur directeur $\vec{f}^{(p)}(t_0)$ est la **tangente** en M_{t_0} à la courbe. Son équation est :

$$(x - x(t_0)) \times y^{(p)}(t_0) - (y - y(t_0)) \times x^{(p)}(t_0) = 0.$$

Ensuite, soient $q > p$ et $\vec{f}^{(q)}(t_0)$ le premier vecteur dérivé non colinéaire à $\vec{f}^{(p)}(t_0)$.

- Si p est impair et q est pair, alors M_{t_0} est un **point ordinaire**.
- Si p et q sont impairs, alors M_{t_0} est un **point d'inflexion**.

- Si p est pair et q est impair, alors M_{t_0} est un **point de rebroussement de première espèce**.
- Si p et q sont pairs, alors M_{t_0} est un **point de rebroussement de seconde espèce**.



Exemples de courbes planes

2. Étude générale

Définition 63.1

La courbe présente une **branche infinie** en $t_0 \in \mathbb{R} \cup \{\pm\infty\}$ si l'une des coordonnées du point $M_t = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$ tend vers $\pm\infty$ quand t tend vers t_0 .

Nous distinguons plusieurs types de branches infinies, énumérés ci-dessous.

Définition 63.2

- Si $\lim_{t \rightarrow t_0} x(t) = \pm\infty$ et $\lim_{t \rightarrow t_0} y(t) = y_0 \in \mathbb{R}$, alors la courbe a une **asymptote horizontale** d'équation $y = y_0$.
- Si $\lim_{t \rightarrow t_0} x(t) = x_0 \in \mathbb{R}$ et $\lim_{t \rightarrow t_0} y(t) = \pm\infty$, alors la courbe a une **asymptote verticale** d'équation $x = x_0$.
- Si $\lim_{t \rightarrow t_0} x(t) = \pm\infty$ et $\lim_{t \rightarrow t_0} y(t) = \pm\infty$, considérons $a = \lim_{t \rightarrow t_0} \frac{y(t)}{x(t)}$.
 - Si $a \in \mathbb{R}^*$, deux cas se présentent. Nous étudions la différence $y(t) - ax(t)$.
 - Si $\lim_{t \rightarrow t_0} y(t) - ax(t) = \pm\infty$, alors la courbe présente une **branche parabolique** de direction $y = ax$.
 - Si $\lim_{t \rightarrow t_0} y(t) - ax(t) = b \in \mathbb{R}$, alors la droite d'équation $y = ax + b$ est une **asymptote oblique**.
 - Si $a = 0$, alors la courbe présente une **branche parabolique** dirigée par l'axe Ox .
 - Si $a = \pm\infty$, alors la courbe présente une **branche parabolique** dirigée par l'axe Oy .

Un plan d'étude d'une courbe plane paramétrée d'équation $f(t) = x(t)\vec{i} + y(t)\vec{j}$ est proposé ci-dessous.

- Nous déterminons le domaine de définition $U \subset \mathbb{R}$ de la fonction $\vec{f}$.
- Nous cherchons ensuite à restreindre le domaine d'étude de la fonction $\vec{f}$.
 - Si la fonction $\vec{f}$ est périodique de période $T \in \mathbb{R}^*$, nous étudions la courbe sur un sous-ensemble $[\alpha; \alpha + T] \subset U$ où $\alpha \in \mathbb{R}$.
 - Nous étudions aussi si les fonctions $t \mapsto x(t)$ et $t \mapsto y(t)$ sont paires ou impaires.

$t \mapsto x(t)$	$t \mapsto y(t)$	Propriété de la courbe
paire	impaire	symétrique par rapport à l'axe Ox
impaire	paire	symétrique par rapport à l'axe Oy
impaire	impaire	symétrique par rapport à l'origine O

Si $x(-t) = y(t)$, $\forall t \in U$, alors la courbe est symétrique par rapport à la première bissectrice d'équation $y = x$.

- Nous calculons les dérivées et tangentes.
- Nous calculons les différentes limites et branches infinies.
- Nous traçons la courbe.

Application Le mouvement d'une roue de vélo

1. Un cycliste roule en ligne droite sur une route. Le mouvement de la valve d'une roue du vélo par rapport au sol lorsque la roue tourne et avance est une cycloïde d'équation :

$$\vec{f}(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} R(t - \sin(t)) \\ R(1 - \cos(t)) \end{pmatrix}$$

où $t \in \mathbb{R}$ et R est un paramètre réel strictement positif. Étudier la courbe paramétrée.

2. Étudier la courbe paramétrée définie par les équations $x(t) = \frac{e^t}{t+1}$ et $y(t) = \frac{te^t}{t+1}$.

1. Domaine de définition

La fonction $\vec{f}$ est définie sur $\mathbb{R}$.

Restriction du domaine de définition

Nous avons pour tout $t \in \mathbb{R}$: $x(t + 2\pi) = x(t) + 2\pi R$ et $y(t + 2\pi) = y(t)$. Nous étudions donc la courbe sur l'intervalle $[-\pi; \pi]$. La totalité de la courbe définie sur $\mathbb{R}$ sera ensuite obtenue en effectuant des translations de vecteurs $2k\pi R \vec{i}$, pour $k \in \mathbb{Z}$, de la courbe restreinte à $[-\pi; \pi]$.

De plus, la fonction $t \mapsto x(t)$ est impaire et la fonction $t \mapsto y(t)$ est paire, la courbe est donc symétrique par rapport à l'axe Oy. Nous réduisons alors notre étude sur l'intervalle $[0; \pi]$.

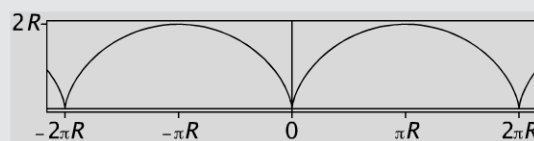
Dérivées et tangentes

La fonction $\vec{f}$ est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$ et nous avons : $x'(t) = R(1 - \cos(t))$ et $y'(t) = R \sin(t)$. Les fonctions x et y sont donc croissantes sur $[0; \pi]$. Plus précisément, x croît de 0 à πR sur $[0; \pi]$ et y croît de 0 à $2R$ sur $[0; \pi]$.

L'origine est un point singulier car $\vec{f}'(0) = \vec{0}$. Comme $\vec{f}''(0) = R\vec{j} \neq \vec{0}$, la tangente en l'origine est l'axe Oy, c'est-à-dire l'axe de symétrie de la courbe.

Tracé

Nous traçons la courbe restreinte à $[0; \pi]$. Par symétrie par rapport à l'axe Oy, nous obtenons la courbe restreinte à $[-\pi; \pi]$. Enfin, par translation de vecteurs $2k\pi R \vec{i}$, $k \in \mathbb{Z}$, nous obtenons la totalité de la courbe définie sur $\mathbb{R}$. Le tracé de la courbe est donné ci-dessous.



Cycloïde

2. Domaine de définition

Le domaine de définition de la courbe paramétrée est $U = \mathbb{R} \setminus \{-1\}$. Il n'y a pas de symétrie.

Dérivées

$$\forall t \in \mathbb{R} \setminus \{-1\}, \quad x'(t) = \frac{te^t}{(t+1)^2}, \quad y'(t) = \frac{e^t(t^2 + t + 1)}{(t+1)^2}.$$

Nous en déduisons qu'il n'y a pas de point singulier car les fonctions x' et y' ne s'annulent jamais simultanément. La fonction x est décroissante de 0 à $-\infty$ sur $]-\infty; -1[$, décroissante de $+\infty$ à 1 sur $] -1; 0]$ et croissante de 1 à $+\infty$ sur $[0; +\infty[$. De même, la fonction y est croissante de 0 à $+\infty$ sur $] -\infty; -1[$ et croissante de $-\infty$ à $+\infty$ sur $] -1; +\infty[$.

Limites et branches infinies

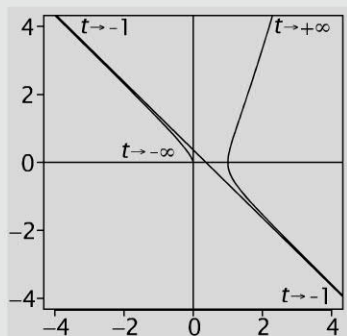
Remarquons que $\lim_{t \rightarrow -\infty} x(t) = \lim_{t \rightarrow -\infty} y(t) = 0$.

Nous avons des branches infinies pour $t \rightarrow -1$ et $t \rightarrow +\infty$. Pour les étudier, nous évaluons les limites $a = \lim_{t \rightarrow -1} \frac{y(t)}{x(t)}$ et $\tilde{a} = \lim_{t \rightarrow +\infty} \frac{y(t)}{x(t)}$.

- Pour $t \rightarrow -1$, comme $a = \lim_{t \rightarrow -1} \frac{y(t)}{x(t)} = -1$, et $\lim_{t \rightarrow -1} y(t) - ax(t) = \lim_{t \rightarrow -1} e^t = e^{-1}$, nous avons une asymptote oblique d'équation $y = -x + e^{-1}$.
- Pour $t \rightarrow +\infty$, comme $\tilde{a} = \lim_{t \rightarrow +\infty} \frac{y(t)}{x(t)} = +\infty$, nous avons une branche parabolique dans la direction Oy .

Tracé

Les tracés de la courbe et de l'asymptote oblique sont donnés ci-dessous.



Pour s'entraîner (solutions p. 536)

63.1 Tracer la courbe d'équations : $x(t) = 3 \cos(t) - 2 \sin^3(t)$, $y(t) = \cos(4t)$.

Même question avec $x(t) = 2 \sin(t) + \cos(t)$, $y(t) = \sin^3(t) + 2 \cos^3(t)$.

63.2 Donner les équations paramétriques d'un cercle de centre O et de rayon $r > 0$.
Même question pour une ellipse de centre O de grand axe $2a > 0$ et de petit axe $2b > 0$.

Courbe plane définie par son équation polaire

Certaines courbes planes peuvent être décrites par un paramètre angulaire, ce sont les courbes planes définies par une équation dite « polaire ».



1. Équation et repère polaires

Définition 64.1

Soit le plan muni du repère orthonormé $(O, \vec{i}, \vec{j})$. Pour tout $\theta \in \mathbb{R}$, nous notons

$$\begin{aligned}\vec{u}(\theta) &= \cos \theta \vec{i} + \sin \theta \vec{j} \\ \vec{v}(\theta) &= \vec{u}\left(\theta + \frac{\pi}{2}\right) = -\sin \theta \vec{i} + \cos \theta \vec{j}.\end{aligned}$$

Le repère orthonormé $(O, \vec{u}(\theta), \vec{v}(\theta))$ est appelé **repère polaire**. Nous avons, pour tout $\theta \in \mathbb{R}$,

$$\frac{d\vec{u}(\theta)}{d\theta} = \vec{v}(\theta) \text{ et } \frac{d\vec{v}(\theta)}{d\theta} = -\vec{u}(\theta).$$

Soient I un intervalle de $\mathbb{R}$ et r une fonction définie sur I à valeurs dans $\mathbb{R}$. Lorsque θ décrit I , l'ensemble des points $M(\theta) = \begin{pmatrix} r(\theta) \cos \theta \\ r(\theta) \sin \theta \end{pmatrix}$, c'est-à-dire tels que $\overrightarrow{OM}(\theta) = r(\theta) \vec{u}(\theta)$ décrit une **courbe plane définie par ses coordonnées polaires**. Nous parlons aussi de **courbe définie par son équation polaire** $\theta \mapsto r(\theta)$.



Une courbe définie en coordonnées polaires est un cas particulier de courbe paramétrée en coordonnées cartésiennes, puisque c'est l'ensemble des points $M(\theta)$ du plan tels que $x(\theta) = r(\theta) \cos \theta$, $y(\theta) = r(\theta) \sin \theta$. C'est donc la courbe associée à la fonction $\vec{f} : I \rightarrow \mathbb{R}^2$ définie par $\vec{f}(\theta) = r(\theta) \vec{u}(\theta)$.

Théorème 64.1

Soient I un intervalle de $\mathbb{R}$, et Γ une courbe définie par son équation polaire $\theta \mapsto r(\theta)$, de classe $\mathcal{C}^1$ sur I . Notons $M(\theta) = \begin{pmatrix} r(\theta) \cos \theta \\ r(\theta) \sin \theta \end{pmatrix}$ un point de la courbe Γ , et notons $\vec{f}(\theta) = \overrightarrow{OM}(\theta) = r(\theta) \vec{u}(\theta)$. Le vecteur

$$\vec{f}'(\theta) = \frac{d\overrightarrow{OM}(\theta)}{d\theta} = r'(\theta) \vec{u}(\theta) + r(\theta) \vec{v}(\theta)$$

est le **vecteur vitesse** du point $M(\theta)$ se déplaçant sur la courbe Γ .

1. Le point $M(\theta)$ de la courbe est dit **régulier** si $r(\theta) \neq 0$ et $r'(\theta) \neq 0$. Un vecteur directeur de la **tangente** en ce point est donné par le vecteur vitesse.

Notons $V = \left(\vec{u}(\theta), \frac{d\overrightarrow{OM}(\theta)}{d\theta} \right)$ l'angle entre la courbe et sa tangente, au point $M(\theta)$.

Nous avons $\tan V = \frac{r(\theta)}{r'(\theta)}$.

- Si $r'(\theta) = 0$, la tangente en $M(\theta)$ est dirigée par $\vec{v}(\theta)$, et $V = \pm \frac{\pi}{2}$.

- Si $r(\theta) = 0$, la tangente en $M(\theta)$ passe par l'origine O et est dirigée par $\vec{u}(\theta)$, et $V = 0$.
 - Si r est de classe $\mathcal{C}^2$ sur I et si $M(\theta)$ est un **point d'inflexion** alors $r^2(\theta) + 2r'(\theta) - r(\theta)r''(\theta) = 0$.
2. Le point $M(\theta)$ est un **point stationnaire** si son vecteur vitesse est nul, c'est-à-dire si $r(\theta) = r'(\theta) = 0$. Les points stationnaires sont nécessairement à l'origine.

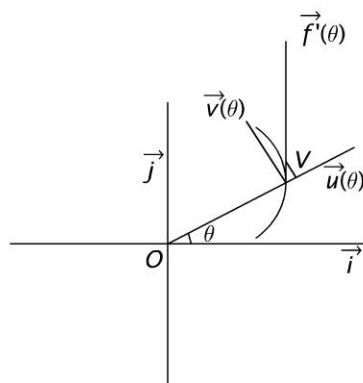


Figure 64.1 – Repère polaire $(O, \vec{u}(\theta), \vec{v}(\theta))$ et tangente en un point régulier de la courbe.

2. Symétries

Indiquons maintenant les propriétés de **symétrie** d'une courbe définie par son équation polaire.

Théorème 64.2

Soit une courbe définie par son équation polaire $\theta \mapsto r(\theta)$, définie sur l'intervalle I .

1. Supposons que la fonction r est **périodique** de période $T \in \mathbb{R}$: pour tout $\theta \in I$, nous avons $\theta - T \in I$, $\theta + T \in I$, et $r(\theta + T) = r(\theta)$. Alors la courbe est invariante par la rotation de centre O et d'angle T . Par exemple, si $T = \pi$, la courbe est invariante par la symétrie centrale de centre O .
2. Supposons qu'il existe un angle réel θ_0 tel que :
 - pour tout $\theta \in I$ tel que $\theta_0 - \theta \in I$, nous avons :

$$r(\theta_0 - \theta) = r(\theta)$$

alors la courbe est symétrique par rapport à l'axe dirigé par le vecteur $\vec{u}\left(\frac{\theta_0}{2}\right)$.

- pour tout $\theta \in I$ tel que $\theta_0 - \theta \in I$, nous avons :

$$r(\theta_0 - \theta) = -r(\theta)$$

alors la courbe est symétrique par rapport à l'axe dirigé par le vecteur $\vec{v}\left(\frac{\theta_0}{2}\right)$.

Exemple

La courbe illustrée à la Figure 64.2 est symétrique par rapport à $\vec{v}\left(\frac{\pi}{4}\right)$, qui est la seconde bissectrice.

3. Asymptotes

Définition 64.2

La courbe définie par son équation polaire présente une **branche infinie** en $\theta_0 \in \mathbb{R} \cup \{\pm\infty\}$ si $\lim_{\theta \rightarrow \theta_0} r(\theta) = \pm\infty$.

Théorème 64.3

Soit une courbe définie par son équation polaire $\theta \mapsto r(\theta)$, définie sur l'intervalle I .

1. Considérons tout d'abord le cas où $\theta \rightarrow \pm\infty$. Deux cas se présentent.
 - Si $\lim_{\theta \rightarrow \pm\infty} r(\theta) = r \in \mathbb{R}$, alors la courbe admet un **cercle asymptotique** de rayon $|r|$. En particulier, si $\lim_{\theta \rightarrow \pm\infty} r(\theta) = 0$, la courbe admet un **point asymptotique**, qui est l'origine.
 - Si $\lim_{\theta \rightarrow \pm\infty} r(\theta) = \pm\infty$, alors la courbe admet une **branche infinie spirale**.
2. Soit maintenant un angle réel θ_0 tel que $\lim_{\theta \rightarrow \theta_0} r(\theta) = \pm\infty$. Considérons le repère $(0, \vec{u}(\theta_0), \vec{v}(\theta_0))$. Dans ce repère, le point $M(\theta)$ de la courbe a pour coordonnées

$$X(\theta) = r(\theta) \cos(\theta - \theta_0), \quad Y(\theta) = r(\theta) \sin(\theta - \theta_0).$$

Nous avons alors $\lim_{\theta \rightarrow \theta_0} X(\theta) = \pm\infty$. Pour $Y(\theta)$, deux cas se présentent.

- Si $\lim_{\theta \rightarrow \theta_0} Y(\theta) = \ell \in \mathbb{R}$, alors la courbe admet une **asymptote** d'équation $Y = \ell$ dans le repère $(O, \vec{u}(\theta_0), \vec{v}(\theta_0))$.
- Si $\lim_{\theta \rightarrow \theta_0} Y(\theta) = \pm\infty$, alors la courbe admet une **branche parabolique** de direction $\vec{u}(\theta_0)$.

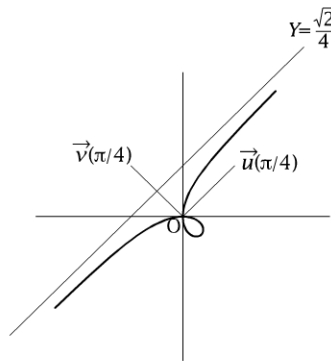


Figure 64.2 – Courbe d'équation $r(\theta) = \frac{\sin \theta \cos \theta}{\sin \theta - \cos \theta}$, admettant une asymptote d'équation $Y = \frac{\sqrt{2}}{4}$ dans le repère $(O, \vec{u}(\frac{\pi}{4}), \vec{v}(\frac{\pi}{4}))$.

Application

La puissance rayonnée et l'étude d'une courbe

1. Une onde électromagnétique divergente se propage dans le vide, et transporte avec elle l'énergie associée aux champs électrique et magnétique. La puissance surfacique moyenne transportée par cette onde est définie par $P_s(\theta) = P_{\max} \sin^2 \theta$, où $\theta \in \mathbb{R}$, et $P_{\max}$ est un réel positif. Étudier la courbe dont l'équation polaire est donnée par P_s .
2. Étudier la courbe définie par son équation polaire $r(\theta) = \cos(2\theta)$, $\theta \in \mathbb{R}$.

1. Nous posons $r(\theta) = P_{\max} \sin^2 \theta$. La fonction r est définie sur $\mathbb{R}$, périodique, de période π . La courbe est donc invariante par une symétrie centrale de centre O . De plus, nous avons $\forall \theta \in \mathbb{R}, r(-\theta) = r(\theta)$: la courbe est donc symétrique par rapport à l'axe Ox . Enfin, nous avons $\forall \theta \in \mathbb{R}, r(\pi - \theta) = P_{\max} \sin^2(\pi - \theta) = r(\theta)$: la courbe est donc symétrique par rapport à l'axe Oy . Il suffit donc d'étudier la courbe sur l'intervalle $\left[0; \frac{\pi}{2}\right]$.

$\forall x \in \left[0; \frac{\pi}{2}\right], r'(\theta) = 2P_{\max} \cos \theta \sin \theta = P_{\max} \sin(2\theta)$. Donc, la fonction r croît de 0 à 1 sur $\left[0; \frac{\pi}{2}\right]$.

Il n'y a pas de branche infinie.

Cherchons les points de la courbe pour lesquels la tangente est perpendiculaire à la courbe. D'après le Théorème 64.1, ce sont les points $M(\theta)$ tels que $r'(\theta) = 0 \Leftrightarrow \sin(2\theta) = 0 \Leftrightarrow \theta = \frac{\pi}{2}$ modulo π .

Le tracé de la puissance en fonction de θ est illustré dans le graphique de gauche de la Figure 64.3.

2. La fonction r est périodique, de période π , donc la courbe est invariante par une symétrie centrale de centre O . La fonction r est paire, donc la courbe est symétrique par rapport à l'axe Ox . De plus, $r\left(\frac{\pi}{2} - \theta\right) = -r(\theta)$ car $\cos(\pi - 2\theta) = -\cos(2\theta)$. La courbe est donc symétrique par rapport à la droite d'équation $y = -x$. Il suffit donc d'étudier la courbe sur l'intervalle $\left[0; \frac{\pi}{4}\right]$.

$\forall x \in \left[0; \frac{\pi}{4}\right], r'(\theta) = -2 \sin(2\theta)$. La fonction r décroît donc de 1 à 0 sur $\left[0; \frac{\pi}{4}\right]$.

Remarquons que nous avons : $\tan(V) = \frac{r(\theta)}{r'(\theta)} = \frac{-\cos(2\theta)}{2 \sin(2\theta)}$, où $V = \left(\vec{u}(\theta), \widehat{\frac{d\vec{OM}(\theta)}{d\theta}}\right)$ désigne l'angle entre la courbe et la tangente au point $M(\theta)$ (comme illustré à la Figure 64.1).

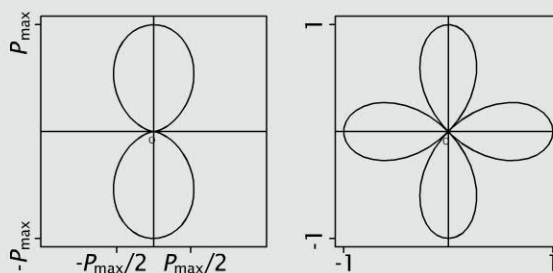


Figure 64.3 – Courbe de l'Application Puissance rayonnée (à gauche), courbe de l'Application Étude d'une courbe (à droite).

Pour s'entraîner (solutions p. 537)

64.1 Tracer la courbe d'équation polaire
 $r(\theta) = \sin \theta + \cos \frac{\theta}{2}$.

64.2 Tracer la courbe d'équation polaire
 $r(\theta) = 1 + 2 \cos \frac{3\theta}{2}$.

64.3 Tracer la courbe d'équation polaire
 $r(\theta) = \frac{1}{\cos \theta + \cos 2\theta}$.

Intégrales doubles et multiples

1. Intégrale sur un pavé de $\mathbb{R}^n$

Soient P un pavé compact de $\mathbb{R}^n$, $P = [a_1; b_1] \times \cdots \times [a_n; b_n] \subset \mathbb{R}^n$, avec $a_1 \leq b_1, \dots, a_n \leq b_n$. Le **volume** de P , noté $\text{vol}(P)$, est égal à $\prod_{j=1}^n |b_j - a_j|$.

Définition 65.1

Soit $(A_i)_{i=1, \dots, k}$ des pavés compacts contenus dans P , nous appelons **fonction en escalier sur P** , une fonction f de la forme, pour tout $\vec{x} \in P$, $f(\vec{x}) = \sum_i^k \lambda_i \mathbb{1}_{A_i}(\vec{x})$, où $\mathbb{1}_{A_i}(\vec{x}) = 1$ si $\vec{x} \in A_i$, 0 sinon. Nous appelons **intégrale sur P de la fonction f en escalier sur P** , notée $\int_P f(\vec{x}) d\vec{x}$, le nombre $\int_P f(\vec{x}) d\vec{x} = \sum_i^k \lambda_i \text{vol}(A_i)$.

Définition 65.2

Nous appelons f **fonction intégrable sur P** une fonction bornée sur P , à valeurs réelles, et telle que, pour tout $\epsilon > 0$, il existe des fonctions en escaliers ϕ et ψ définies sur P telles que $\phi \leq f \leq \psi$ et $\int_P (\psi(\vec{x}) - \phi(\vec{x})) d\vec{x} < \epsilon$. Si f est intégrable sur P , alors $\sup_{\phi \leq f} \int_P \phi(\vec{x}) d\vec{x} = \inf_{f \leq \psi} \int_P \psi(\vec{x}) d\vec{x}$. Cette valeur est appelée **intégrale de f sur P** et est notée $\int_P f(\vec{x}) d\vec{x}$ ou $\int_P f(x_1, \dots, x_n) dx_1 \dots dx_n$.

Théorème 65.1 Théorème de Fubini

Soit f une fonction définie et continue sur $[a; b] \times [c; d] \subset \mathbb{R}^2$ et à valeurs réelles. Les fonctions $x \mapsto \int_c^d f(x, y) dy$ et $y \mapsto \int_a^b f(x, y) dx$ sont continues respectivement sur $[a; b]$ et $[c; d]$ et nous avons :

$$\int_{[a; b] \times [c; d]} f(x, y) dx dy = \int_a^b \left(\int_c^d f(x, y) dy \right) dx = \int_c^d \left(\int_a^b f(x, y) dx \right) dy.$$

Corollaire 65.1

Soient $g : [a; b] \rightarrow \mathbb{R}$ et $h : [c; d] \rightarrow \mathbb{R}$ deux fonctions continues sur des intervalles de $\mathbb{R}$, alors nous avons :

$$\int_a^b \int_c^d g(x) h(y) dy dx = \left(\int_a^b g(x) dx \right) \times \left(\int_c^d h(y) dy \right).$$

2. Intégrale sur un domaine plan

Définition 65.3

- Un **domaine plan** D est dit **de type I** s'il est délimité par les graphes de deux fonctions continues de x , c'est-à-dire :

$$D = \{(x, y) \in \mathbb{R}^2, \text{ tel que } a \leq x \leq b, g_1(x) \leq y \leq g_2(x)\}$$

où g_1 et g_2 sont deux fonctions continues sur un intervalle $[a; b]$ de $\mathbb{R}$ et à valeurs réelles.

- Un **domaine plan** D est dit de **type II** s'il est délimité par les graphes de deux fonctions continues de y , c'est-à-dire :

$$D = \{(x, y) \in \mathbb{R}^2, \text{ tel que } c \leq y \leq d, h_1(y) \leq x \leq h_2(y)\}$$

où h_1 et h_2 sont deux fonctions continues sur un intervalle $[c; d]$ de $\mathbb{R}$ et à valeurs réelles.

Théorème 65.2

Soient $f : D \rightarrow \mathbb{R}$ une fonction continue à valeurs réelles et D un domaine de type I ou de type II. Si f est à valeurs strictement positives, $\int \int_D f(x, y) dx dy$ s'interprète comme le volume V du solide S ayant pour base D et se dressant jusqu'à la surface $z = f(x, y)$.

- Si D un domaine plan de type I, alors :

$$\int \int_D f(x, y) dx dy = \int_a^b \left(\int_{g_1(x)}^{g_2(x)} f(x, y) dy \right) dx.$$

- Si D un domaine plan de type II, alors :

$$\int \int_D f(x, y) dx dy = \int_c^d \left(\int_{h_1(y)}^{h_2(y)} f(x, y) dx \right) dy.$$

Théorème 65.3 Linéarité de l'intégrale

Soient $\alpha, \beta \in \mathbb{R}$, D un domaine plan de type I ou II, et deux fonctions continues $f : D \rightarrow \mathbb{R}$ et $g : D \rightarrow \mathbb{R}$. Nous avons :

$$\int \int_D (\alpha f(x, y) + \beta g(x, y)) dx dy = \alpha \int \int_D f(x, y) dx dy + \beta \int \int_D g(x, y) dx dy.$$

Nous rappelons ci-dessous comment calculer l'aire d'une surface paramétrée à l'aide d'une intégrale double.

 **Fiche 62**

Théorème 65.4

Soient D un domaine plan de type I ou II de $\mathbb{R}^2$ et S une surface paramétrée définie par : $\vec{r}(u, v) = x(u, v) \vec{i} + y(u, v) \vec{j} + z(u, v) \vec{k}$, où $(u, v) \in D$ et $\vec{r}$ est une fonction de classe $\mathcal{C}^1$ sur D . Si S est couverte une seule fois lorsque (u, v) parcourt le domaine D , alors l'aire de la surface S est le nombre :

$$A(S) = \int \int_D \|\vec{r}'_u \wedge \vec{r}'_v\| dx dy, \quad \text{où } \vec{r}'_u = \frac{\partial x}{\partial u} \vec{i} + \frac{\partial y}{\partial u} \vec{j} + \frac{\partial z}{\partial u} \vec{k}, \quad \text{et } \vec{r}'_v = \frac{\partial x}{\partial v} \vec{i} + \frac{\partial y}{\partial v} \vec{j} + \frac{\partial z}{\partial v} \vec{k}.$$

Théorème 65.5 Changement de variables

Soient R et S deux ouverts de $\mathbb{R}^2$, et $\varphi : S \rightarrow R$ une fonction telle que $\forall (u, v) \in S$, $\varphi(u, v) = (x, y) \in R$, un $\mathcal{C}^1$ -difféomorphisme de S sur R dont le jacobien est borné sur S . Alors pour toute fonction $f : R \rightarrow \mathbb{R}$ continue et bornée sur R ,

$$\int \int_R f(x, y) dx dy = \int \int_S f(x(u, v), y(u, v)) |J(u, v)| du dv,$$

où $|J(u, v)| = \left| \frac{\partial x}{\partial u} \frac{\partial y}{\partial v} - \frac{\partial x}{\partial v} \frac{\partial y}{\partial u} \right|$ est la valeur absolue du jacobien de la transformation $\varphi : (u, v) \mapsto (x(u, v), y(u, v))$.

Toutes les définitions et tous les théorèmes précédents s'étendent aux **intégrales triples** et **multiplées**. Exprimons les théorèmes de Fubini et de changement de variables dans ce cas général.

Théorème 65.6 Théorème de Fubini

Soit $P = [a_1; b_1] \times \dots \times [a_n; b_n]$ un pavé de $\mathbb{R}^n$. Soit $f : P \rightarrow \mathbb{R}$ une fonction continue, alors :

$$\int_P f(x_1, \dots, x_n) dx_1 \dots dx_n = \int_{a_1}^{b_1} \left(\int_{a_2}^{b_2} \left(\dots \left(\int_{a_n}^{b_n} f(x_1, \dots, x_n) dx_n \right) \dots \right) dx_2 \right) dx_1.$$

Théorème 65.7 Changement de variables

Soient R et S deux ouverts de $\mathbb{R}^n$, et $\phi : S \rightarrow R$ une fonction telle que $\forall \vec{u} \in S, \phi(\vec{u}) = \vec{x} \in R$, un $\mathcal{C}^1$ -difféomorphisme de S sur R dont le jacobien est borné sur S . Alors pour toute fonction $f : R \rightarrow \mathbb{R}$ continue et bornée sur R , nous avons :

$$\int_R f(x_1, \dots, x_n) dx_1 \dots dx_n = \int_S f(u_1, \dots, u_n) |J(\vec{u})| du_1 \dots du_n,$$

où $J(\vec{u})$ est le jacobien de la transformation $\phi : \vec{u} \mapsto \vec{x}$.

Application

1. Aire d'une sphère

Calculer l'aire d'une sphère de rayon R , de représentation paramétrique : $x = R \sin \phi \cos \theta, y = R \sin \phi \sin \theta, z = R \cos \phi$ avec $(\phi, \theta) \in D = [0; \pi] \times [0; 2\pi]$.

2. Intégrale double

Calculer l'intégrale $I = \iint_D e^{\frac{x+y}{x-y}} dx dy$ où D est la région de sommets : $(2; 0), (4; 0), (0; -2), (0; -4)$.

3. Intégrale triple

Calculer l'intégrale triple $I = \iiint_D \left(\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \right)^\alpha dx dy dz$ avec $a > 0, b > 0, c > 0, \alpha > 0$ et $D = \left\{ (x, y, z) \in \mathbb{R}^3 \text{ tel que } \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} < 1 \right\}$.

1. Notons $\vec{r}(\phi, \theta) = x(\phi, \theta) \vec{i} + y(\phi, \theta) \vec{j} + z(\phi, \theta) \vec{k}$. Le produit vectoriel des vecteurs tangents est :

$$\vec{r}'_\phi \wedge \vec{r}'_\theta = \begin{vmatrix} R \cos \phi \sin \theta & -R \sin \phi \\ R \sin \phi \cos \theta & 0 \end{vmatrix} \vec{i} - \begin{vmatrix} R \cos \phi \cos \theta & -R \sin \phi \\ -R \sin \phi \sin \theta & 0 \end{vmatrix} \vec{j} + \begin{vmatrix} R \cos \phi \cos \theta & R \cos \phi \sin \theta \\ -R \sin \phi \sin \theta & R \sin \phi \cos \theta \end{vmatrix} \vec{k}$$

Donc $\vec{r}'_\phi \wedge \vec{r}'_\theta = R^2 \sin^2(\phi) \cos(\theta) \vec{i} + R^2 \sin^2(\phi) \sin(\theta) \vec{j} + R^2 \sin(\phi) \cos(\phi) \vec{k}$.

Ainsi : $\|\vec{r}'_\phi \wedge \vec{r}'_\theta\| = R^2 \sin(\phi)$ car $\sin(\phi) \geq 0$. Nous avons alors :

$$A = \int \int_D \|\vec{r}'_\phi \wedge \vec{r}'_\theta\| d\phi d\theta = \int_0^{2\pi} \int_0^\pi R^2 \sin(\phi) d\phi d\theta = 4\pi R^2.$$

2. Nous faisons le changement de variables : $u = x + y$, $v = x - y$.

Donc, $x = \frac{1}{2}(u + v)$, $y = \frac{1}{2}(u - v)$. Le changement de variables $(x, y) = \varphi(u, v)$ défini sur $S = \{(u, v), 2 \leq v \leq 4, -v \leq u \leq v\}$ et à valeurs dans D est bijectif car nous pouvons exprimer (u, v) en fonction de (x, y) et vice-versa, il est de classe $\mathcal{C}^1$ sur S et son jacobien ne s'annule jamais sur S :

$$J(u, v) = \begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{vmatrix} = -\frac{1}{2}.$$

$$\text{Nous avons alors } I = \int_2^4 \left(\int_{-v}^v e^{\frac{u}{v}} \frac{1}{2} du \right) dv = \frac{1}{2} \int_2^4 v(e - e^{-1}) dv = 3(e - e^{-1}).$$

3. Nous effectuons le changement de variables suivant :

$$x = a\rho \cos \theta \sin \phi, y = b\rho \sin \theta \sin \phi, z = c\rho \cos \phi \text{ avec } (\rho, \phi, \theta) \in]0; 1[\times]0; \pi[\times]0; 2\pi[.$$

Le jacobien de la transformation est $J(\rho, \phi, \theta) = \rho^2 abc \sin \phi$.

$$\text{Donc } I = \int_0^1 \int_0^\pi \int_0^{2\pi} \rho^{2\alpha} abc \rho^2 \sin \phi d\rho d\phi d\theta = \frac{4abc\pi}{2\alpha + 3}. \text{ Si } \alpha = 0, \text{ c'est le volume d'un ellipsoïde.}$$

Pour s'entraîner (solutions p. 537)

65.1 Intégrale double

Soit $D = [0; 7] \times [3; 5]$. Calculer l'intégrale double : $I = \int \int_D (x^2 - 7y^3) dx dy$.

65.2 Volume d'un solide

Calculer le volume du solide qui se dresse sur le domaine D du plan Oxy délimité par la droite $y = 3x$ et la parabole $y = 2x^2$ et couverte par la parabolôide

$$z = 3x^2 + 2y^2. \text{ Nous donnons : } D = \left\{ (x, y) \in \mathbb{R}^2, 0 \leq x \leq \frac{3}{2}, 2x^2 \leq y \leq 3x \right\}.$$

65.3 Calculer $I = \int \int_D (x + 3y) dx dy$ sur le demi-disque unité se trouvant dans le demi-plan $x \geq 0$.

65.4 Calculer $I = \int \int_D (x + ye^x) dx dy$ sur le triangle de sommets $(0; 1)$, $(0; -2)$, $(2; 3)$.

Sommabilité et intégrales doubles ou multiples généralisées

1. Séries indexées par une partie de $\mathbb{Z}$

Définition 66.1

Soient I une partie de $\mathbb{N}$ ou de $\mathbb{Z}$ et $(x_i)_{i \in I}$ une famille de nombres réels. Pour définir la **convergence absolue de la somme** $\sum_{i \in I} x_i$ quatre cas sont à distinguer : la famille I peut être finie, minorée, majorée ou n'être ni minorée, ni majorée.

1. Si I est une partie finie de $\mathbb{Z}$ de cardinal k , $I = \{i_1, \dots, i_k\}$, alors la somme $\sum_{i \in I} x_i = \sum_{k=1}^n x_{i_k}$ converge absolument.
2. Si I est une partie minorée de $\mathbb{Z}$, alors il existe $n_0 \in \mathbb{Z}$ tel que $I \subset \llbracket n_0; +\infty \llbracket$ et $I = \{i_k; k \in \mathbb{N}\}$, la somme $\sum_{i \in I} x_i$ converge absolument si la série $\sum_{k=0}^{+\infty} x_{i_k}$ converge absolument.
3. Si I est une partie majorée de $\mathbb{Z}$, alors $J = -I = \{-i; i \in I\}$ est une partie minorée de $\mathbb{Z}$. La somme $\sum_{i \in I} x_i$ converge absolument si la série $\sum_{j \in J} x_{-j}$ converge absolument.
4. Si I est une partie de $\mathbb{Z}$ ni majorée ni minorée, alors posons $I^- = \{i; i \in I \text{ et } i < 0\}$ et $I^+ = \{i; i \in I \text{ et } i \geq 0\}$. I^- est une partie de $\mathbb{Z}$ majorée et I^+ est une partie de $\mathbb{Z}$ minorée. La somme $\sum_{i \in I} x_i$ converge absolument si les séries $\sum_{i \in I^-} x_i$ et $\sum_{i \in I^+} x_i$ convergent absolument.

Une famille $(x_i)_{i \in I}$ de nombres réels est **sommable**, c'est-à-dire la somme $\sum_{i \in I} x_i$ existe, si la somme $\sum_{i \in I} x_i$ converge absolument.

 Fiche 46

Proposition 66.1

Soient $I \subset \mathbb{N}$ ou $\subset \mathbb{Z}$ et $(x_i)_{i \in I}$ une famille sommable de nombres réels.

1. Pour toute partie $J \subset I$, la somme $\sum_{j \in J} x_j$ converge absolument et $\sum_{j \in J} |x_j| \leq \sum_{i \in I} |x_i|$.
2. Soient $J \subset I$ et $K \subset \mathbb{N}$. Si $(J_k)_{k \in K}$ est une partition de J , alors $\sum_{j \in J} x_j = \sum_{k \in K} \sum_{j_k \in J_k} x_{j_k}$.

 Fiche 3

2. Séries doubles

Définition 66.2

Soit $(u_{i,j})_{(i,j) \in I \times J}$ une famille de nombres réels positifs. Si pour tout $i \in I$, la somme $\sum_{j \in J} u_{i,j}$ converge et est égale à a_i et si la somme $\sum_{i \in I} a_i$ converge et a pour somme U , nous appelons $\sum_{(i,j) \in I \times J} u_{i,j}$ **série double convergente** de somme U ou $(u_{i,j})_{(i,j) \in I \times J}$ famille sommable.

Proposition 66.2

Si $\sum_{(i,j) \in I \times J} u_{i,j}$ est une série double convergente de somme U , alors pour tout $j \in J$, $\sum_{i \in I} u_{i,j}$ converge et est égale à b_j et $\sum_{j \in J} b_j$ converge et a pour somme U . Nous avons ainsi en cas de convergence $\sum_{(i,j) \in I \times J} u_{i,j} = \sum_{i \in I} (\sum_{j \in J} u_{i,j}) = \sum_{j \in J} (\sum_{i \in I} u_{i,j})$.

Proposition 66.3

Soient $(u_{i,j})_{(i,j) \in I \times J}$ et $(v_{i,j})_{(i,j) \in I \times J}$ deux familles de nombres réels tels que pour tout $(i, j) \in I \times J$, $0 \leq v_{i,j} \leq u_{i,j}$. Si $\sum_{(i,j) \in I \times J} u_{i,j}$ est une série double convergente, alors il en va de même de $\sum_{(i,j) \in I \times J} v_{i,j}$ et $0 \leq \sum_{(i,j) \in I \times J} v_{i,j} \leq \sum_{(i,j) \in I \times J} u_{i,j}$.

Définition 66.3

Soit $(u_{i,j})_{(i,j) \in I \times J}$ une famille de nombres réels. Nous appelons $\sum_{(i,j) \in I \times J} u_{i,j}$ **série double absolument convergente**, lorsque la série double $\sum_{(i,j) \in I \times J} |u_{i,j}|$ est convergente. Dans ce cas $(u_{i,j})_{(i,j) \in I \times J}$ est également appelée famille sommable.

Pour tout $(i, j) \in I \times J$ posons $u_{i,j}^+ = \max(u_{i,j}, 0)$ et $u_{i,j}^- = \max(-u_{i,j}, 0)$. La somme d'une série double $\sum_{(i,j) \in I \times J} u_{i,j}$ absolument convergente est $\sum_{(i,j) \in I \times J} u_{i,j} = \sum_{(i,j) \in I \times J} u_{i,j}^+ - \sum_{(i,j) \in I \times J} u_{i,j}^-$.

Théorème 66.1 Théorème de Fubini

Si $\sum_{(i,j) \in I \times J} u_{i,j}$ est une série double absolument convergente alors :

1. pour tout $i \in I$, $\sum_{j \in J} u_{i,j}$ converge et $\sum_{i \in I} \sum_{j \in J} u_{i,j}$ converge ;
2. pour tout $j \in J$, $\sum_{i \in I} u_{i,j}$ converge et $\sum_{j \in J} \sum_{i \in I} u_{i,j}$ converge ;
3. les égalités suivantes sont vérifiées
$$\sum_{(i,j) \in I \times J} u_{i,j} = \sum_{i \in I} \left(\sum_{j \in J} u_{i,j} \right) = \sum_{j \in J} \left(\sum_{i \in I} u_{i,j} \right).$$

Théorème 66.2

Si $\sum_{(i,j) \in I \times J} u_{i,j}$ est une série double absolument convergente, alors pour toute partition $(A_k)_{k \in K}$ de $I \times J$:

1. pour tout $k \in K$, la série double $\sum_{(i,j) \in A_k} u_{i,j}$ converge absolument ;
2. la série $\sum_{k \in K} \left(\sum_{(i,j) \in A_k} u_{i,j} \right)$ converge absolument ;
3. l'égalité suivante est vérifiée :
$$\sum_{(i,j) \in I \times J} u_{i,j} = \sum_{k \in K} \left(\sum_{(i,j) \in A_k} u_{i,j} \right).$$

3. Intégrales doubles et multiples généralisées sur $\mathbb{R}^n$ **Définition 66.4**

Soit f une fonction définie sur $\mathbb{R}^n$, à valeurs réelles positives et intégrable sur tout pavé compact de $\mathbb{R}^n$. f est appelée **fonction intégrable sur $\mathbb{R}^n$** si la borne supérieure des intégrales de f sur tout pavé compact de $\mathbb{R}^n$ est finie. Dans ce cas, l'intégrale de f sur $\mathbb{R}^n$ est définie par :

$$\int_{\mathbb{R}^n} f(\vec{x}) d\vec{x} = \sup_{P \text{ pavé compact } \subset \mathbb{R}^n} \int_P f(\vec{x}) d\vec{x}.$$

Une fonction f définie sur $\mathbb{R}^n$ et à valeurs réelles est **intégrable sur $\mathbb{R}^n$** si $|f|$ l'est.

Proposition 66.4

Soit $(A_k)_{k \in \mathbb{N}}$ une suite croissante d'ensembles bornés dont les fonctions caractéristiques $\mathbb{1}_{A_k}$, où $\mathbb{1}_{A_k}(\vec{x}) = 1$ si $\vec{x} \in A_k$, 0 sinon, sont intégrables et telles que $\bigcup_{k \in \mathbb{N}} A_k = \mathbb{R}^n$. f est intégrable sur $\mathbb{R}^n$ si et seulement si, pour tout $k \in \mathbb{N}$, f est intégrable sur A_k et la suite des intégrales $\int_{A_k} f(\vec{x}) d\vec{x}$ est bornée.

Théorème 66.3 Fubini

Soit f une fonction à valeurs réelles, définie et intégrable sur $\mathbb{R}^2$. Si pour tout x réel la fonction $y \mapsto f(x, y)$ est intégrable sur $\mathbb{R}$ et si la fonction

$$x \mapsto \int_{-\infty}^{+\infty} |f(x, y)| dy$$

est intégrable sur $\mathbb{R}$, alors nous avons :

$$\int_{\mathbb{R}^2} f(x, y) dx dy = \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} f(x, y) dy \right) dx.$$

Proposition 66.5 Intversion des intégrales

Soit f une fonction à valeurs réelles, définie et intégrable sur $\mathbb{R}^2$. Si pour tout x réel la fonction $y \mapsto f(x, y)$ est intégrable sur $\mathbb{R}$ et pour tout y réel la fonction $x \mapsto f(x, y)$ est intégrable sur $\mathbb{R}$ et si les fonctions

$$x \mapsto \int_{-\infty}^{+\infty} |f(x, y)| dy \text{ et } y \mapsto \int_{-\infty}^{+\infty} |f(x, y)| dx$$

sont intégrables sur $\mathbb{R}$, alors nous avons :

$$\int_{\mathbb{R}^2} f(x, y) dx dy = \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} f(x, y) dy \right) dx = \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} f(x, y) dx \right) dy.$$

Application

1. Soit $\lambda \in \mathbb{R}$. La famille de nombre réels $\left(\frac{\lambda^{n+m}}{n!m!} \right)_{(n,m) \in \mathbb{N}^2}$ est-elle sommable ?

2. Soient $0 < a < b$. En utilisant une formule de Fubini, calculer :

$$I = \int_0^{+\infty} \frac{e^{-ax} - e^{-bx}}{x} dx$$

1. Pour tout réel λ , la série $\sum_{m=0}^{+\infty} \frac{\lambda^m}{m!}$ converge et sa somme vaut $\exp(\lambda)$. Soit $a_n = \sum_{m=0}^{+\infty} \frac{\lambda^n \lambda^m}{n! m!} = \frac{\lambda^n}{n!} \sum_{m=0}^{+\infty} \frac{\lambda^m}{m!} = \frac{\exp(\lambda) \lambda^n}{n!}$.

La série $\sum_{n=0}^{+\infty} a_n$ converge et sa somme est égale à $\exp(2\lambda)$.

Ainsi la famille $\left(\frac{\lambda^{n+m}}{n!m!} \right)_{(n,m) \in \mathbb{N}^2}$ est sommable. Nous en déduisons que

$$\sum_{(n,m) \in \mathbb{N}^2} \frac{\lambda^{n+m}}{n!m!} = \sum_{n=0}^{+\infty} \sum_{m=0}^{+\infty} \frac{\lambda^{n+m}}{n!m!} = \sum_{m=0}^{+\infty} \sum_{n=0}^{+\infty} \frac{\lambda^{n+m}}{n!m!} = \exp(2\lambda).$$



Commençons par remarquer que : $\frac{e^{-ax} - e^{-bx}}{x} = \int_a^b e^{-xy} dy = \int_{\mathbb{R}} e^{-xy} \mathbb{1}_{[a;b]}(y) dy$.

2. Soient $f : (x, y) \in \mathbb{R}^2 \mapsto e^{-xy} \mathbb{1}_{[a;b]}(y) \mathbb{1}_{[0;+\infty[}(x)$ et (A_n) la suite de parties de $\mathbb{R}^2 : A_n = [-n; n] \times [-n; n]$. Sur chaque A_n , la fonction f est positive, bornée et est le produit d'une fonction continue sur $\mathbb{R}^2$ par une fonction en escalier sur $\mathbb{R}^2$ donc intégrable sur A_n . Le théorème de Fubini s'applique et ainsi $0 \leq \int_{A_n} f(x, y) dx dy = \int_{-n}^n \int_{-n}^n e^{-xy} \mathbb{1}_{[a;b]}(y) \mathbb{1}_{[0;+\infty[}(x) dx dy = \int_{-n}^n \mathbb{1}_{[a;b]}(y) \int_0^n e^{-xy} dx dy = \int_a^{\min(b;n)} \frac{1 - e^{-n}}{y} dy = (1 - e^{-n}) [\ln y]_a^{\min(b;n)} = (1 - e^{-n}) \ln \left(\frac{\min(b;n)}{a} \right) \leq \ln \left(\frac{b}{a} \right)$.

La fonction $f : (x, y) \in \mathbb{R}^2 \mapsto e^{-xy} \mathbb{1}_{[a;b]}(y)$ est donc intégrable sur $\mathbb{R}^2$. Puis chacune des applications $x \mapsto f(x, y)$ et $y \mapsto f(x, y)$ sont intégrables sur $\mathbb{R}$. Enfin, $y \mapsto \int_{\mathbb{R}} f(x, y) dx$ et $x \mapsto \int_{\mathbb{R}} f(x, y) dy$ sont intégrables sur $\mathbb{R}$.

Le théorème de Fubini s'applique et nous avons alors :

$$\begin{aligned} I &= \int_0^{+\infty} \frac{e^{-ax} - e^{-bx}}{x} dx = \int_{\mathbb{R}} \mathbb{1}_{[0;+\infty[}(x) \int_{\mathbb{R}} e^{-xy} \mathbb{1}_{[a;b]}(y) dy dx = \int_{\mathbb{R}} \int_{\mathbb{R}} f(x, y) dy dx \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} f(x, y) dx dy = \int_{\mathbb{R}} \mathbb{1}_{[a;b]}(y) \int_0^{+\infty} e^{-xy} dx dy = \int_{-n}^n \mathbb{1}_{[a;b]}(y) \frac{1}{y} dy = [\ln y]_a^b = \ln \left(\frac{b}{a} \right) \end{aligned}$$

Pour s'entraîner (solutions p. 538)

66.1 Séries de nombres premiers

Pour tout $n \in \mathbb{N}$, $n \geq 2$, soit q_n le plus grand facteur premier de n . Si $k \in \mathbb{N}^*$, soit $\mathcal{E}_k = \{n \in \mathbb{N}^*, n \geq 2 \mid q_n = p_k\}$, où $(p_k)_{k \geq 1}$ désigne la suite strictement croissante des nombres premiers qui commence pour $p_1 = 2$. Soit enfin $M_k = \prod_{i=1}^k \frac{1}{1 - \frac{1}{p_i}}$.

1. Montrer que, pour tout $k \in \mathbb{N}^*$, la famille $\left(\frac{1}{n q_n} \right)_{n \in \mathcal{E}_k}$ est sommable de somme $\frac{M_k}{p_k^2}$.

2. Montrer que la série $\sum_{n=2}^{+\infty} \frac{1}{n q_n}$ converge.

66.2 Montrer que l'intégrale double généralisée, $I = \int \int_D \frac{dx dy}{x + y^2}$, converge sur le domaine $D = \{(x, y) \in \mathbb{R}^2 : 0 \leq x \leq 1, y^2 \leq x\}$.

66.3 Calculer

$$I = \int_{[0;+\infty[\times [0;+\infty[} 4xy e^{-x^2-y^2} dx dy.$$

66.4 1. Étudier la nature de l'intégrale double généralisée $\int \int_{\mathbb{R}^2} \frac{dx dy}{(1 + x^2 + y^2)^\alpha}$ selon la valeur du nombre réel α .

2. Montrer que l'intégrale double généralisée $\int \int_{\mathbb{R}^2} \frac{\sin(x+y) dx dy}{(1 + x^2 + y^2)^2}$ est absolument convergente.

66.5 1. Calculer l'intégrale double généralisée $K = \int \int_{\mathbb{R}^+ \times \mathbb{R}^+} e^{-x^2-y^2} dx dy$, en utilisant la suite de compacts $(K_n = \{(x, y) \mid x \geq 0, y \geq 0, x^2 + y^2 \leq n^2\})$.

2. En déduire, en utilisant la suite de compacts $(L_n = \{(x, y) \mid [0; n] \times [0; n]\})$ et le théorème de Fubini, la valeur de l'intégrale simple généralisée $\int_0^{+\infty} e^{-x^2} dx$.

66.6 Montrer que l'intégrale généralisée $\int \int_{\mathbb{R}^2} \sin(x^2 + y^2) dx dy$ est divergente en considérant les pavés $P_n = [-n; n] \times [-n; n]$ et les disques $D_n = \{(x, y) \mid x^2 + y^2 \leq n^2\}$.

Dans cette Fiche, l'intégration porte sur une courbe C . Nous considérons ici des courbes C planes ou dans l'espace $\mathbb{R}^3$ définies par leur équation vectorielle : $\vec{r}(t) = x(t)\vec{i} + y(t)\vec{j}$ si la courbe est plane, ou $\vec{r}(t) = x(t)\vec{i} + y(t)\vec{j} + z(t)\vec{k}$ si la courbe est dans l'espace $\mathbb{R}^3$, pour $t \in I = [a; b]$ intervalle de $\mathbb{R}$. Les courbes considérées sont lisses, c'est-à-dire que la fonction $\vec{r}$ est de classe $\mathcal{C}^1$ sur $]a; b[$ et $\vec{r}'(t) \neq \vec{0}$ en tout point de $]a; b[$.

1. Intégrale par rapport à l'abscisse curviligne

Nous rappelons que la longueur de C est définie par :

$$L = \int_a^b \|\vec{r}'(t)\| dt$$

ce que nous notons aussi

$$L = \int_C ds.$$

Définition 67.1

Soit f une fonction définie et bornée sur la courbe C . L'**intégrale de f sur C par rapport à l'abscisse curviligne**, notée $\int_C f ds$ est définie par

$$\int_C f ds = \int_a^b f(\vec{r}(t)) \times \|\vec{r}'(t)\| dt = \int_a^b f(x(t), y(t), z(t)) \times \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2} dt.$$



Si f est la fonction Identité, alors nous retrouvons la longueur de L .

2. Intégrale curviligne

Définition 67.2

Soient une courbe lisse C définie par une fonction vectorielle $\vec{r} :]a; b[\rightarrow \mathbb{R}^3$, et $\vec{T}(t) = \frac{\vec{r}'(t)}{\|\vec{r}'(t)\|}$ le vecteur tangent unitaire au point $\vec{r}(t)$ de la courbe C . Soit $\vec{F}$ un champ vectoriel continu défini sur la courbe C . L'**intégrale curviligne** de $\vec{F}$ le long de C est :

$$\int_C \vec{F} \cdot d\vec{r} = \int_a^b \vec{F}(\vec{r}(t)) \cdot \vec{r}'(t) dt = \int_C \vec{F} \cdot \vec{T} ds,$$

où $ds = \|\vec{r}'(t)\| dt$.

Exemple

Soit C l'hélice circulaire de rayon $r > 0$ et de hauteur $h\pi/2$, avec $h > 0$, paramétrée par $\vec{r}(t) = r \cos t \vec{i} + r \sin t \vec{j} + ht \vec{k}$ pour $t \in [0; \frac{\pi}{2}]$. Soit le champ vectoriel $\vec{F}$ défini sur $\mathbb{R}^3$ par $\vec{F}(x, y, z) = (y, z, x)$. Alors $\int_C \vec{F} \cdot d\vec{r} = \frac{\pi r}{4}(2h - r)$.



- L'intégrale curviligne est aussi appelée **travail** ou **circulation** du champ de vecteurs $\vec{F}$ sur la courbe C . Le mot *travail* évoque plutôt l'idée que $\vec{F}$ est une force, et le mot *circulation* celle que $\vec{F}$ est la vitesse d'un élément fluide.
- Pour calculer une intégrale curviligne, il faut définir un paramétrage du chemin d'intégration.

3. Gradient et fonction différentiable**Théorème 67.1 Circulation d'un champ de gradient**

Soit C une courbe lisse donnée par la fonction vectorielle $\vec{r}$ définie sur l'intervalle $[a; b]$, soit f une fonction numérique différentiable dont le vecteur gradient $\vec{\nabla} f$ est continu sur U un ouvert **connexe par arcs** de $\mathbb{R}^3$ contenant la courbe C , alors :

Fiches 58 et 59

$$\int_C \vec{\nabla} f \cdot d\vec{r} = f(\vec{r}(b)) - f(\vec{r}(a)).$$



- La conséquence est que l'intégrale curviligne d'un gradient continu sur un ouvert connexe par arcs est indépendante du chemin.
- f est parfois appelé **potentiel**.
- Dans les conditions du théorème, l'intégrale curviligne dépend seulement des valeurs du potentiel f en $\vec{r}(a)$ et $\vec{r}(b)$, elle ne dépend pas du chemin C pris pour aller de $\vec{r}(a)$ à $\vec{r}(b)$, du moment que C reste dans le domaine connexe par arcs U .
- L'intégrale curviligne d'un gradient sur une courbe fermée d'un ouvert connexe par arcs est nulle !

En fait, les gradients sont les seuls champs de vecteurs ayant cette propriété que leur intégrale curviligne est indépendante du chemin. C'est l'objet du théorème suivant.

Théorème 67.2

Soit U un ouvert **connexe par arcs** de $\mathbb{R}^3$, et $\vec{F}$ un champ de vecteurs continu défini sur U à valeurs dans $\mathbb{R}^3$. Nous avons les équivalences suivantes :

- $\vec{F}$ dérive d'un potentiel f , fonction numérique définie sur U (i.e. $\vec{F} = \vec{\nabla} f$),
- Pour toute courbe fermée C de U , $\int_C \vec{F} \cdot d\vec{r} = 0$,
- L'intégrale curviligne de $\vec{F}$ entre deux points $\vec{x}$ et $\vec{y}$ de U est indépendante de la courbe C de U considérée pour aller de $\vec{x}$ à $\vec{y}$.

Si en plus, $\vec{F}$ est de classe $\mathcal{C}^1$ et si U est **convexe** alors i), ii), iii) sont équivalents avec le fait que la matrice jacobienne de $\vec{F}$ est symétrique, ou encore que le rotationnel de $\vec{F}$ est nul.

Fiche 60



Si $\vec{F}$ vérifie l'un des points i), ii), iii), nous disons que $\vec{F}$ est un **champ conservatif**.

En particulier, si nous nous plaçons dans le plan $\mathbb{R}^2$, nous avons le théorème suivant.

Théorème 67.3

Si $\vec{F}$ est un champ vectoriel conservatif de classe $\mathcal{C}^1$ sur un ouvert connexe par arcs U du plan, tel que $\forall (x, y) \in U, \vec{F}(x, y) = P(x, y) \vec{i} + Q(x, y) \vec{j}$, alors nous avons :

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}.$$

4. Théorème de Green dans le plan

Théorème 67.4 Théorème de Green dans le plan

Soit C une courbe plane, fermée, lisse par morceaux, et soit D le domaine du plan délimité par C . Soient P et Q deux fonctions définies et de classe $\mathcal{C}^1$ sur un ouvert qui contient D , alors :

$$\int_C (Pdx + Qdy) = \int \int_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy.$$

Application

Une intégrale curviligne – Quelques propriétés d'un fil

1. Calculer $I = \int_C (2y^2 dx + 5xy dy)$ où C est la frontière du demi-anneau circulaire supérieur formé par les cercles : $x^2 + y^2 = 1$ et $x^2 + y^2 = 9$.

2. Soit C un fil constitué d'un matériau de densité $\varphi : C \rightarrow \mathbb{R}$, où $\varphi(x, y, z) = x^2 + y^2 + z^2$.

La masse M du fil est définie par $M = \int_C \varphi(x, y, z) ds$, son centre de masse est défini par le

point $(\bar{x}, \bar{y}, \bar{z})$ tel que $\bar{x} = \frac{1}{M} \int_C x \varphi(x, y, z) ds$, $\bar{y} = \frac{1}{M} \int_C y \varphi(x, y, z) ds$, $\bar{z} = \frac{1}{M} \int_C z \varphi(x, y, z) ds$,

et le moment d'inertie du fil par rapport à l'axe Oz est défini par $I_z = \int_C (x^2 + y^2) \varphi(x, y, z) ds$.

Calculer la masse M , le centre de masse $\bar{z}$, et le moment d'inertie I_z d'un enroulement d'une bobine ayant une forme hélicoïdale paramétrée par $\vec{r}(t) = r \cos t \vec{i} + r \sin t \vec{j} + ht \vec{k}$, où $r, h > 0$, et $t \in [0; 2\pi]$.

1. La courbe C est dans le plan, lisse par morceaux. Nous notons les fonctions $P(x, y) = 2y^2$ et $Q(x, y) = 5xy$. Les fonctions P et Q sont de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$. Soit D l'intérieur de la courbe C , nous avons

$$D = \{(x, y) \in \mathbb{R}^2, \text{ tel que } x = r \cos \theta, y = r \sin \theta, \text{ avec } 1 \leq r \leq 3, 0 \leq \theta \leq \pi\}.$$

D'après le théorème de Green dans le plan, nous avons

$$I = \int \int_D \left(\frac{\partial(5xy)}{\partial x} - \frac{\partial(2y^2)}{\partial y} \right) dx dy = \int \int_D y dx dy.$$

Nous posons $x = r \cos \theta, y = r \sin \theta$, donc en passant en coordonnées polaires, $I =$

$\int_0^\pi \int_1^3 r \sin \theta |J| dr d\theta$, où $|J|$ est la valeur absolue du jacobien :

$$J = \begin{vmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{vmatrix} = r.$$



Fiche 60

Donc, $|J| = r$. Nous avons alors : $I = \int_0^\pi \sin \theta d\theta \int_1^3 r^2 dr = [-\cos \theta]_0^\pi \left[\frac{r^3}{3} \right]_1^3 = \frac{52}{3}$.

2. La courbe est lisse et pour tout $t \in [0; 2\pi]$, nous avons $\|\vec{r}'(t)\| = \sqrt{r^2 + h^2}$. La masse du fil est égale à $M = \int_0^{2\pi} \varphi(r \cos t, r \sin t, ht) \times \sqrt{r^2 + h^2} dt = \sqrt{r^2 + h^2} \int_0^{2\pi} (r^2 + h^2 t^2) dt = 2\pi \sqrt{r^2 + h^2} \left(r^2 + \frac{4\pi^2}{3} h^2 \right)$.

De même, $\int_C z \varphi(x, y, z) ds = \int_0^{2\pi} ht \varphi(r \cos t, r \sin t, ht) \times \sqrt{r^2 + h^2} dt = h \sqrt{r^2 + h^2} \int_0^{2\pi} t (r^2 + h^2 t^2) dt = 2\pi h \sqrt{r^2 + h^2} \left(\pi r^2 + 2\pi^3 h^2 \right)$. D'où $\bar{z} = h \frac{\pi r^2 + 2\pi^3 h^2}{r^2 + \frac{4\pi^2}{3} h^2}$.

Enfin, le moment d'inertie du fil par rapport à l'axe Oz , est égal à $I_z = \int_C (x^2 + y^2) \times (x^2 + y^2 + z^2) ds = r^2 \sqrt{r^2 + h^2} \int_0^{2\pi} (r^2 + h^2 t^2) dt = r^2 \sqrt{r^2 + h^2} \left(2\pi r^2 + \frac{8\pi^3}{3} h^2 \right)$.

Pour s'entraîner (solutions p. 539)

67.1 Calculer $I = \int_C (x^5 dx + x^2 y dy)$ où C est la courbe triangulaire composée des trois segments de $(0; 0)$ à $(2; 0)$, de $(2; 0)$ à $(0; 3)$, de $(0; 3)$ à $(0; 0)$.

67.2 Calculer le travail effectué par le champ de forces $\vec{F}(x, y) = x \vec{i} - x^2 y \vec{j}$ pour déplacer un point matériel le long du demi-cercle $\vec{r}(t) = \cos(t) \vec{i} + \sin(t) \vec{j}$, $0 \leq t \leq \pi$.

67.3 Déterminer si le champ vectoriel : $\vec{F}(x, y) = (2x + y) \vec{i} + (3x - y) \vec{j}$ est conservatif.

67.4 Calculer $I = \int_C (2 + x^2 y) ds$ où C est le demi-cercle supérieur de rayon 1 d'équation $x^2 + y^2 = 1$.

67.5 Calculer $I = \int_C 2x ds$ où C se compose de l'arc C_1 de la parabole $y = x^2$ entre $(0; 0)$ et $(1; 1)$ suivi du segment vertical C_2 qui relie $(1; 1)$ à $(1; 2)$.

67.6 Calculer $I = \int_C y^2 dx + 3xy dy$, où C est la frontière du demi-anneau circulaire supérieur formé par les cercles $x^2 + y^2 = 1$ et $x^2 + y^2 = 4$.

67.7 Soit le champ de vecteurs $\vec{F}(x, y) = P(x, y) \vec{i} + Q(x, y) \vec{j}$ avec $P(x, y) = \frac{2x}{y}$ et $Q(x, y) = \frac{1 - x^2}{y^2}$. Soit $U = \{(x, y) \in \mathbb{R}^2, y > 0\}$ le demi-plan supérieur ouvert, et C une courbe d'extrémités $\vec{a} = (a_1, a_2)$, $\vec{b} = (b_1, b_2)$ incluse dans U . Calculer $\int_C \vec{F} \cdot d\vec{r}$.

67.8 Calculer l'intégrale $\int_C \frac{xdy - ydx}{x^2 + y^2}$ dans les cas suivants :

1. si C est une courbe incluse dans le demi-plan supérieur ouvert $\{(x, y) \in \mathbb{R}^2, y > 0\}$, d'extrémités $\vec{a}$ et $\vec{b}$,
2. si C est le cercle trigonométrique.

Soit S une surface paramétrée par le champ vectoriel $\vec{r}(u, v) = x(u, v)\vec{i} + y(u, v)\vec{j} + z(u, v)\vec{k}$, pour $(u, v) \in D$, où D est un ouvert connexe par arcs du plan $\mathbb{R}^2$. Si le champ $\vec{r}$ est de classe $\mathcal{C}^1$ sur D , nous définissons un vecteur unitaire normal à S , en tout point $\vec{r}(u, v)$ régulier par

 Fiche 62

$$\vec{n}(u, v) = \frac{\frac{\partial \vec{r}(u, v)}{\partial u} \wedge \frac{\partial \vec{r}(u, v)}{\partial v}}{\left\| \frac{\partial \vec{r}(u, v)}{\partial u} \wedge \frac{\partial \vec{r}(u, v)}{\partial v} \right\|}.$$

1. Intégrale de surface

Définition 68.1

Soit une surface S décrite par le champ vectoriel $\vec{r}(u, v) = x(u, v)\vec{i} + y(u, v)\vec{j} + z(u, v)\vec{k}$, pour $(u, v) \in D$, où D est un ouvert connexe par arcs du plan $\mathbb{R}^2$. La fonction $\vec{r}$ est supposée de classe $\mathcal{C}^1$ sur D .

Soit f une fonction définie sur la surface S et à valeurs réelles. Nous définissons l'intégrale de surface de f sur S par

$$\iint_S f(x, y, z) dS = \iint_D f(\vec{r}(u, v)) \times \left\| \frac{\partial \vec{r}(u, v)}{\partial u} \wedge \frac{\partial \vec{r}(u, v)}{\partial v} \right\| du dv.$$



- En particulier, si $f = 1$, nous obtenons l'aire de la surface S :

$$\iint_S dS = \iint_D \left\| \frac{\partial \vec{r}(u, v)}{\partial u} \wedge \frac{\partial \vec{r}(u, v)}{\partial v} \right\| du dv.$$

- Toute surface S d'équation $\vec{z} = g(x, y)$, avec $(x, y) \in D$ peut être vue comme une surface paramétrée par $\vec{r}(x, y) = x\vec{i} + y\vec{j} + g(x, y)\vec{k}$. Nous avons alors :

$$\iint_S f(x, y, z) dS = \iint_D f(x, y, g(x, y)) \sqrt{\left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2 + 1} dx dy.$$

 Fiche 62

Exemple

Calculons l'intégrale de surface $I = \iint_S x^2 dS$ où S est la sphère représentée paramétriquement par : $\vec{r}(\phi, \theta) = \sin \phi \cos \theta \vec{i} + \sin \phi \sin \theta \vec{j} + \cos \phi \vec{k}$, avec $(\phi, \theta) \in D =]0; \pi[\times]0; 2\pi[$.

Donc, $\iint_S x^2 dS = \int_0^{2\pi} \int_0^\pi (\sin \phi \cos \theta)^2 \left\| \frac{\partial \vec{r}}{\partial \phi} \wedge \frac{\partial \vec{r}}{\partial \theta} \right\| d\phi d\theta.$

Donc, $I = \int_0^{2\pi} \cos^2 \theta d\theta \int_0^\pi \sin^3 \phi d\phi = \left[\frac{\theta}{2} + \frac{\sin(2\theta)}{4} \right]_0^{2\pi} \times \left[-\cos \phi + \frac{\cos^3 \phi}{3} \right]_0^\pi = \frac{4\pi}{3}.$

2. Flux d'un champ vectoriel à travers une surface

Définition 68.2

Soit S une surface paramétrée par le champ vectoriel $\vec{r}(u, v) = x(u, v)\vec{i} + y(u, v)\vec{j} + z(u, v)\vec{k}$, pour $(u, v) \in D$, où D est un ouvert connexe par arcs du plan $\mathbb{R}^2$. Nous supposons que $\vec{r}$ est de classe $\mathcal{C}^1$ sur D , et en tout point régulier de S , nous notons $\vec{n}(u, v)$ un vecteur unitaire normal à S . Soit $\vec{F} : S \rightarrow \mathbb{R}^3$ un champ vectoriel défini sur S . Le **flux de $\vec{F}$ à travers la surface S** est défini par :

$$\iint_S \vec{F} \cdot d\vec{S} = \iint_S \vec{F} \cdot \vec{n} dS = \iint_D \vec{F}(\vec{r}(u, v)) \cdot \left(\frac{\partial \vec{r}(u, v)}{\partial u} \wedge \frac{\partial \vec{r}(u, v)}{\partial v} \right) du dv.$$



Soit une surface S d'équation $z = g(x, y)$, avec $(x, y) \in D$. La surface S est alors paramétrée par $\vec{r}(x, y) = x\vec{i} + y\vec{j} + g(x, y)\vec{k}$.

Si le champ vectoriel $\vec{F}$ s'écrit $\vec{F}(x, y) = P(x, y)\vec{i} + Q(x, y)\vec{j} + R(x, y)\vec{k}$, alors

$$\iint_S \vec{F} \cdot d\vec{S} = \iint_D \left(-P \frac{\partial g}{\partial x} - Q \frac{\partial g}{\partial y} + R \right) dx dy,$$

$$\text{car } \vec{F} \cdot \left(\frac{\partial \vec{r}}{\partial x} \wedge \frac{\partial \vec{r}}{\partial y} \right) = (P\vec{i} + Q\vec{j} + R\vec{k}) \cdot \left(-\frac{\partial g}{\partial x}\vec{i} - \frac{\partial g}{\partial y}\vec{j} + \vec{k} \right).$$



Dans la définition précédente, l'orientation de la normale à la surface, $\vec{n}$, est importante.

Si nous prenons la normale orientée dans le sens contraire, nous avons $-\vec{n} = \frac{\partial \vec{r}(u, v)}{\partial v} \wedge \frac{\partial \vec{r}(u, v)}{\partial u}$. Dans le calcul du flux de $\vec{F}$, nous obtenons alors le signe contraire.

Exemple

Soit un fluide constitué d'un ensemble de particules. Chaque particule notée par ses coordonnées (x, y, z) possède une vitesse $\vec{V}(x, y, z)$. C'est le champ de vitesse du flot. Nous supposons que ce champ $\vec{V}$ ne dépend que de la position de la particule et non du temps. Désignons par $\rho(x, y, z)$ la densité (masse par unité de volume) du fluide en la particule (x, y, z) . Si le fluide est incompressible, la densité est constante dans le fluide. Pour un fluide compressible tel qu'un gaz, la densité peut varier dans le fluide. Dans tous les cas, la densité est une fonction numérique associée au mouvement. Le produit de la densité et de la vitesse est noté $\vec{F} : \vec{F}(x, y, z) = \rho(x, y, z)\vec{V}(x, y, z)$. C'est un champ de vecteurs appelé densité de flux du flot. Le vecteur $\vec{F}$ a la même direction que la vitesse et ses dimensions sont $\frac{\text{masse}}{(\text{unité de surface})(\text{unité de temps})}$. En d'autres termes, $\vec{F}$ nous donne la masse de fluide, par unité de surface et de temps, en mouvement dans la direction $\vec{V}$ en tout point (x, y, z) du fluide. Soit S une surface de $\mathbb{R}^3$, alors la masse de fluide en mouvement à travers S par unité de temps est définie par :

$$\iint_S \vec{F} \cdot d\vec{S}.$$

3. Formule de Stokes

Théorème 68.1 Théorème de Stokes

Soit S une surface paramétrée par $\vec{r}$ défini sur un ouvert D de $\mathbb{R}^2$. Notons C le bord de S , supposé fermé. Nous avons $S = \vec{r}(D)$ et $C = \vec{r}(\partial D)$, où

- ∂D est le bord de D , c'est une courbe fermée plane, régulière par morceaux,

- $\vec{r}$ est une application bijective dont les composantes admettent des dérivées secondes continues sur un ouvert contenant $D \cup \partial D$.

Soit ensuite un champ de vecteurs $\vec{F}$ de classe $\mathcal{C}^1$ sur S . Alors nous avons

$$\int_C \vec{F} \cdot d\vec{r} = \iint_S \text{rot} \vec{F} \cdot d\vec{S} = \iint_D \text{rot} \vec{F}(\vec{r}(u, v)) \cdot \left(\frac{\partial \vec{r}(u, v)}{\partial u} \wedge \frac{\partial \vec{r}(u, v)}{\partial v} \right) du dv,$$

où le rotationnel rot a été défini à la Fiche 60.



- Autrement dit, la circulation du champ de vecteurs $\vec{F}$ le long d'un contour fermé C est égal au flux de son rotationnel à travers toute surface ouverte S s'appuyant sur ce contour.
- Nous n'avons pas précisé le sens de la normale à considérer. Tout ceci est lié à des considérations sur les orientations des surfaces. Nous n'entrerons pas dans ces détails dans ce livre.

Définition 68.3

Un champ de vecteurs $\vec{F}$ de classe $\mathcal{C}^1$ sur $\mathbb{R}^3$ est à **circulation conservative** si et seulement si, en tout point de l'espace $\mathbb{R}^3$, nous avons $\text{rot} \vec{F} = \vec{0}$.

Application

La formule de Stokes et le flux de chaleur à travers une gaine

1. Calculer $I = \int_C \vec{F} \cdot d\vec{r}$ où $\vec{F}(x, y, z) = y^2 \vec{i} - x^2 \vec{j} + z \vec{k}$ et où C est la courbe d'intersection du plan : $y + z = 3$ avec le cylindre $x^2 + y^2 = 1$.

2. Soit S une gaine cylindrique de longueur L et de base circulaire de rayon ρ . La gaine est une conduite dans laquelle circule un fluide de température $T(r)$ ne dépendant que de la distance r à l'axe de la conduite. Sur l'axe de la conduite, la température est égale à T_1 , et à l'extérieur de la gaine, elle est égale à T_2 . La température du fluide est donnée par $T(r) = T_1 + (T_2 - T_1) \ln\left(\frac{e-1}{\rho} r + 1\right)$ où $e = \exp(1)$. Calculer la quantité (i.e. le flux) de chaleur s'échappant de la gaine chaque seconde, donnée par l'équation

$$\dot{Q} = -\lambda \iint_S \frac{\partial T}{\partial r} dS,$$

où λ est la conductivité thermique de la gaine.

1. Nous prenons comme surface délimitée par C la région du plan $y + z = 3$ que ceinture le cercle $x^2 + y^2 = 1$. Nous notons S cette surface. La surface S est paramétrée par

$$\vec{r} : \begin{cases} D =]0; 2\pi[\times]0; 1[\longrightarrow \mathbb{R}^3 \\ (\theta, \rho) \longmapsto (\rho \cos \theta, \rho \sin \theta, 3 - \rho \sin \theta) \end{cases}$$

La fonction $\vec{r}$ est de classe $\mathcal{C}^2$ sur D , et $\frac{\partial \vec{r}(\theta, \rho)}{\partial \theta} \wedge \frac{\partial \vec{r}(\theta, \rho)}{\partial \rho} = (0, -\rho, -\rho)$.

La fonction $\vec{F}$ est de classe $\mathcal{C}^1$ sur $\mathbb{R}^3$ et $\text{rot} \vec{F} = -(2x + 2y) \vec{k}$.

Nous avons donc, par le théorème de Stokes,

$$\begin{aligned} I &= \int_C \vec{F} \cdot d\vec{r} = \iint_S \text{rot} \vec{F} \cdot d\vec{S} = \iint_D \text{rot} \vec{F}(\vec{r}(\theta, \rho)) \cdot \left(\frac{\partial \vec{r}(\theta, \rho)}{\partial \theta} \wedge \frac{\partial \vec{r}(\theta, \rho)}{\partial \rho} \right) d\theta d\rho \\ &= \int_0^{2\pi} \int_0^1 (2\rho \cos \theta + 2\rho \sin \theta) \rho d\theta d\rho = \int_0^{2\pi} \left(\frac{2 \cos \theta}{3} + \frac{2 \sin \theta}{3} \right) d\theta = 0. \end{aligned}$$

2. La gaine est paramétrée par le champ $\vec{r} : (\theta, z) \mapsto (\rho \cos \theta, \rho \sin \theta, z)$ défini sur $D =]0; 2\pi[\times]0; L[$. La fonction $\vec{r}$ est de classe $\mathcal{C}^1$ sur D , et $\left\| \frac{\partial \vec{r}(\theta, z)}{\partial \theta} \wedge \frac{\partial \vec{r}(\theta, z)}{\partial z} \right\| = \|(\rho \cos \theta, \rho \sin \theta, 0)\| = \rho$.

De plus, $\frac{\partial T}{\partial r} = \frac{(T_2 - T_1) \times (e - 1)}{(e - 1) \times r + \rho}$, et sur la surface S , c'est-à-dire sur la gaine, cette quantité vaut $\frac{(T_2 - T_1) \times (e - 1)}{e\rho}$.

La quantité de chaleur s'échappant chaque seconde de la gaine est donc

$$\dot{Q} = -\lambda \int_0^L \int_0^{2\pi} \frac{(T_2 - T_1) \times (e - 1)}{e\rho} \rho d\theta dz = -2\pi\lambda L(T_2 - T_1) \times \frac{(e - 1)}{e}.$$

Nous n'avions en fait pas besoin de paramétrer S . En effet, ici les calculs sont simples car

$$\begin{aligned} \dot{Q} &= -\lambda \iint_S \frac{(T_2 - T_1) \times (e - 1)}{e\rho} dS = -\lambda \frac{(T_2 - T_1) \times (e - 1)}{e\rho} \iint_S dS \\ &= -\lambda \frac{(T_2 - T_1) \times (e - 1)}{e\rho} \times \text{Aire}(S) = -2\pi\lambda L(T_2 - T_1) \times \frac{(e - 1)}{e}. \end{aligned}$$

Pour s'entraîner (solutions p. 540)

68.1 Calculer $\iint_S \vec{F} \cdot d\vec{S}$ où $\vec{F}(x, y, z) = x\vec{i} + y\vec{j} + z\vec{k}$ et où S est la frontière du solide fermé par le paraboloïde $z = 1 - 2x^2 - y^2$ et le plan $z = 0$. Nous considérons la partie supérieure du paraboloïde.

68.2 Calculer $\iint_S \text{rot} \vec{F} \cdot d\vec{S}$ où $\vec{F}(x, y, z) = xz\vec{i} + yz\vec{j} + xy\vec{k}$ et S est la portion de la sphère : $x^2 + y^2 + z^2 = 9$ qui se trouve à l'intérieur du cylindre : $x^2 + y^2 = 1$ et au-dessus du plan Oxy .

68.3 Calculer $I = \iint_S \vec{F} \cdot d\vec{S}$ où $\vec{F}(x, y, z) = y\vec{i} + x\vec{j} + z\vec{k}$ et S est la frontière du solide fermé par le paraboloïde $z = 1 - x^2 - y^2$ et le plan $z = 0$.

68.4 Calculer le flux du champ $\vec{F}$ à travers la surface S , où S est l'hémisphère $x^2 + y^2 + z^2 = 1, z \geq 0$, et $\vec{F}$ le champ de vecteurs $\vec{F}(x, y, z) = x\vec{i} + y\vec{j}$.

1. Définitions

Définitions 69.1

1. Notons $\mathcal{C}$, l'ensemble des applications continues par morceaux de $[0; +\infty[$ dans $\mathbb{C}$. Pour $f \in \mathcal{C}$, notons $\mathcal{D}_f$ l'ensemble des nombres p de $\mathbb{C}$ tels que l'application $t \mapsto e^{-pt} f(t)$ soit intégrable sur $[0; +\infty[$. Soit $\mathcal{F}$ l'ensemble des applications f de $\mathcal{C}$ telles que $\mathcal{D}_f \neq \emptyset$.
2. Soit $\alpha \in \mathbb{R}$. Une fonction f de $\mathcal{C}$ est d'ordre exponentiel α à l'infini si, et seulement si, il existe un nombre réel M tel que : $|f(t)| \leq M e^{\alpha t}$, $\forall t \geq 0$.
3. Pour $f \in \mathcal{F}$, l'application $\mathcal{L} : \mathcal{D}_f \rightarrow \mathbb{C}$ s'appelle la transformée de Laplace de f . Nous écrivons $F(p) = (\mathcal{L}[f])(p)$.

Proposition 69.1

Si une fonction f est d'ordre exponentiel α , alors $\emptyset \neq \{p \in \mathbb{C} \mid \Re(p) > \alpha\} \subset \mathcal{D}_f$ et ainsi $f \in \mathcal{F}$. Ainsi la transformée de Laplace F de f existe pour tout $\Re(p) > \alpha$. Elle est de plus de classe $\mathcal{C}^\infty$ sur $[\alpha; +\infty[$.

Proposition 69.2

Soit $f \in \mathcal{F}$. $\forall p_0 \in \mathcal{D}_f$, $\forall p \in \mathbb{C}$, $(\Re(p) \geq \Re(p_0) \implies p \in \mathcal{D}_f)$.

Définition 69.2

Soit $\sigma_f = \inf_{p \in \mathcal{D}_f} \Re(p)$. σ_f est appelée l'abscisse de convergence de f pour la transformée de Laplace.



Si $\sigma_f \in \mathbb{R}$, alors $\mathcal{D}_f$ est le demi-plan, ouvert ou fermé, limité à gauche par la droite d'équation $\Re(p) = \sigma_f$. Si $\sigma_f = -\infty$, alors $\mathcal{D}_f = \mathbb{C}$.

2. Propriétés de la transformée de Laplace

► Limites

Proposition 69.3

Soit $f \in \mathcal{F}$. Si $\lim_{t \rightarrow +\infty} f(t)$ existe, alors $\lim_{t \rightarrow +\infty} f(t) = \lim_{p \rightarrow 0} pF(p)$. Si $\lim_{t \rightarrow 0^+} f(t)$ existe, alors $\lim_{t \rightarrow 0^+} f(t) = \lim_{p \rightarrow +\infty} pF(p)$.

► Transformée de Laplace inverse

Proposition 69.4

Soient $(f, g) \in \mathcal{F}^2$. Si $\mathcal{L}[f] = \mathcal{L}[g]$, alors pour tout $a > 0$, pour tout $t \in [0; a]$, sauf peut-être en un nombre fini de valeurs, $f(t) = g(t)$. Si f est continue alors $f = g$ sur tout $\mathbb{R}_+$.

Définition 69.3

La fonction f , appelée original, est la transformée de Laplace inverse de la fonction F si pour tout $x \in \mathbb{R}_+$, $f(x) = \mathcal{L}^{-1}[F](x)$.

Proposition 69.5

Soit $a \neq 0$. $\mathcal{L}^{-1}[F(ap)](x) = \frac{1}{a}f\left(\frac{x}{a}\right)$. Soit $a \in \mathbb{R}$, $\mathcal{L}^{-1}[F(p+a)] = e^{-ax}f(x)$. Soit $n \in \mathbb{N}$, $\mathcal{L}^{-1}[F^n(p)] = (-1)^n x^n f(x)$.

► Tableau de synthèse

Les fonctions f ci-dessous sont supposées appartenir à $\mathcal{F}$ et suffisamment régulières.

Propriété	Domaine temporel	Domaine « p »
Linéarité	$af(t) + bg(t)$	$aF(p) + bG(p)$
Dérivée de F	$tf(t)$	$-F'(p)$
Dérivée n -ième de F	$t^n f(t)$	$(-1)^n F^{(n)}(p)$
Dérivée de f f est supposée dérivable et $\mathcal{D}_{f'} \neq \emptyset$.	$f'(t)$	$pF(p) - f(0^-)$
Dérivée seconde de f f est supposée deux fois dérivable et $\mathcal{D}_{f''} \neq \emptyset$.	$f''(t)$	$p^2 F(p) - pf(0^-) - f'(0^-)$
Dérivée n -ième de f f est supposé n -fois dérivable et $\mathcal{D}_{f^{(n)}} \neq \emptyset$.	$f^{(n)}(t)$	$p^n F(p) - p^{n-1}f(0^-) - \dots - f^{(n-1)}(0^-)$
Primitive de F	$\frac{f(t)}{t}$	$\int_p^\infty F(\sigma) d\sigma$
Primitive de f $u(t)$ est la fonction échelon de Heaviside. $(u * f)(t)$ est le produit de convolution de $u(t)$ et $f(t)$.	$\int_0^t f(\tau) d\tau = (u * f)(t)$	$\frac{1}{p} F(p)$
Dilatation du temps	$f(at)$	$\frac{1}{ a } F\left(\frac{p}{a}\right)$
Décalage sur p	$e^{at} f(t)$	$F(p-a)$
Décalage du temps $u(t)$ est la fonction échelon de Heaviside	$f(t-a)u(t-a)$	$e^{-ap} F(p)$
Multiplication L'intégration est effectuée le long de la ligne verticale $\Re(\sigma) = c$ qui est entièrement située dans le rayon de convergence de F .	$f(t)g(t)$	$\frac{1}{2\pi i} \lim_{T \rightarrow \infty} \int_{c-iT}^{c+iT} F(\sigma)G(p-\sigma) d\sigma$
Produit de convolution $f(t)$ et $g(t)$ sont étendues sur $\mathbb{R}$ pour la définition du produit de convolution.	$(f * g)(t) = \int_0^t f(\tau)g(t-\tau) d\tau$	$F(p) \cdot G(p)$
Conjugaison complexe	$\bar{f}(t)$	$\bar{F}(\bar{p})$
Fonction périodique $f(t)$ est une fonction périodique de période T telle que $f(t) = f(t+T)$, $\forall t \geq 0$.	$f(t)$	$\frac{1}{1-e^{-Tp}} \int_0^T e^{-pt} f(t) dt$

3. Transformées usuelles

Le domaine de définition des fonctions f de ce tableau est égal à $\mathbb{R}_+^*$ ou, en présence d'un retard égal à τ , à $[\tau; +\infty[$. De plus, $n \in \mathbb{N}$, $(t, \alpha, \beta, \tau, \omega) \in \mathbb{R}^5$, $q \in]-1; +\infty[$, $p \in \mathbb{C}$.

	Domaine temporel $x(t) = \mathcal{L}^{-1}\{X(p)\}$	Transformée de Laplace $X(p) = \mathcal{L}\{x(t)\}$	Région de convergence
Fonction de Dirac	$\delta(t)$	1	$\forall p$
Délai idéal	$\delta(t - \tau)$	$e^{-\tau p}$	
Puissance n -ième	$\frac{t^n}{n!}$	$\frac{1}{p^{n+1}}$	$\Re(p) > 0$
Puissance q -ième	$\frac{t^q}{\Gamma(q+1)}$	$\frac{1}{p^{q+1}}$	$\Re(p) > 0$
Fonction de Heaviside	$\gamma(t)$	$\frac{1}{p}$	$\Re(p) > 0$
Échelon retardé	$\gamma(t - \tau)$	$\frac{e^{-\tau p}}{p}$	$\Re(p) > 0$
Rampe	t	$\frac{1}{p^2}$	$\Re(p) > 0$
Retard avec décalage	$\frac{(t - \tau)^n}{n!} e^{-\alpha(t - \tau)} \cdot \gamma(t - \tau)$	$\frac{e^{-\tau p}}{(p + \alpha)^{n+1}}$	$\Re(p) > 0$
Décroissance exponentielle	$e^{-\alpha t}$	$\frac{1}{p + \alpha}$	$\Re(p) > -\alpha$
Approche exponentielle	$(1 - e^{-\alpha t})$	$\frac{\alpha}{p(p + \alpha)}$	$\Re(p) > 0$
Sinus	$\sin(\omega t)$	$\frac{\omega}{p^2 + \omega^2}$	$\Re(p) > 0$
Cosinus	$\cos(\omega t)$	$\frac{p}{p^2 + \omega^2}$	$\Re(p) > 0$
Sinus hyperbolique	$\text{sh}(\alpha t)$	$\frac{\alpha}{p^2 - \alpha^2}$	$\Re(p) > \alpha $
Cosinus hyperbolique	$\text{ch}(\alpha t)$	$\frac{p}{p^2 - \alpha^2}$	$\Re(p) > \alpha $
Sinus avec décroissance exponentielle	$e^{-\alpha t} \sin(\omega t)$	$\frac{\omega}{(p + \alpha)^2 + \omega^2}$	$\Re(p) > -\alpha$
Cosinus avec décroissance exponentielle	$e^{-\alpha t} \cos(\omega t)$	$\frac{p + \alpha \omega}{(p + \alpha)^2 + \omega^2}$	$\Re(p) > -\alpha$
Racine n -ième	$\sqrt[n]{t}$	$p^{-(n+1)/n} \cdot \Gamma\left(1 + \frac{1}{n}\right)$	$\Re(p) > 0$
Logarithme	$\ln\left(\frac{t}{t_0}\right)$	$-\frac{t_0}{p} [\ln(t_0 p) + \gamma]$	$\Re(p) > 0$
Fonction de Bessel du premier type, d'ordre n	$J_n(\omega t)$	$\frac{\omega^n (p + \sqrt{p^2 + \omega^2})^{-n}}{\sqrt{p^2 + \omega^2}}$	$\Re(p) > 0$ ($n > -1$)
Fonction de Bessel modifiée du premier type, d'ordre n	$I_n(\omega t)$	$\frac{\omega^n (p + \sqrt{p^2 - \omega^2})^{-n}}{\sqrt{p^2 - \omega^2}}$	$\Re(p) > \omega $
Fonction d'erreur	$\text{erf}(t)$	$\frac{e^{p^2/4} \text{erfc}(p/2)}{p}$	$\Re(p) > 0$

δ est la **fonction généralisée de Dirac** égale à 0 partout sauf en 0 où elle est égale à $+\infty$, γ la **fonction échelon de Heaviside** égale à 1 sur $\mathbb{R}_+$ et nulle sur $\mathbb{R}_-^*$, Γ la **fonction Gamma** d'Euler, γ la **constante d'Euler-Mascheroni** égale à $\lim_{n \rightarrow +\infty} \sum_{k=1}^n \frac{1}{k} - \ln n$, erf la **fonction d'erreur** définie par, $\forall z \in \mathbb{R} : \text{erf}(z) = 2\Phi(z\sqrt{2}) - 1$ où Φ est la fonction de répartition de la loi normale centrée et réduite et erfc le **complémentaire de la fonction d'erreur** erf définie par, $\forall z \in \mathbb{R} : \text{erfc}(z) = 1 - \text{erf}(z)$.

 Fiche 53

 Fiche 83

1. Circuit électrique. Nous appliquons une tension électrique $e(t)$ sur une résistance R , une bobine d'inductance L et un condensateur de capacité C branchés en série dans lesquels circulent un courant d'intensité $i(t)$. Calculer $i(t)$ si nous appliquons sur à ces composants un échelon de tension : $e(t) = E$ pour $t \geq 0$ et $e(t) = 0$ pour $t < 0$.

2. Ressort et force sinusoïdale. Soit un ressort dont une masse est suspendue à l'extrémité, nous supposons que le ressort subit une force sinusoïdale. Nous avons : $\frac{d^2x(t)}{dt^2} + k^2x(t) = A \sin(wt)$ avec $x(0) = x_0$ et $x'(0) = v_0$. Supposons que $k \neq 0$ et $w \neq \pm k$. Trouver $x(t)$ en utilisant la transformée de Laplace.

1. Nous avons $e(t) = L \frac{di}{dt} + Ri + \frac{1}{C} \int i dt$ avec $i(0) = 0$. Donc, $\frac{E}{p} = LpI(p) + RI(p) + \frac{1}{C} \frac{I(p)}{p}$ et nous avons alors $I(p) = \frac{E}{L} \frac{1}{p^2 + \frac{R}{L}p + \frac{1}{LC}}$. Étude de $P(p) = p^2 + \frac{R}{L}p + \frac{1}{LC}$. $\Delta = \frac{R^2}{L^2} - \frac{4}{LC}$.

Si $\Delta > 0$, alors nous avons racines réelles distinctes $p_1 = \frac{-R}{2L} + \frac{\sqrt{\Delta}}{2}$, et $p_2 = \frac{-R}{2L} - \frac{\sqrt{\Delta}}{2}$, ce qui donne $I(p) = \frac{E}{L} \frac{1}{(p - p_1)(p - p_2)} = \frac{E}{L} \left(\frac{A}{p - p_1} - \frac{A}{p - p_2} \right)$

avec $A = \frac{1}{\sqrt{\Delta}}$. Par conséquent, $i(t) = \frac{E}{L\sqrt{\Delta}} \left(e^{\frac{-R}{2L}t} (e^{p_1t} - e^{p_2t}) \right)$, c'est-à-dire, comme

$\text{sh}(x) = \frac{e^x - e^{-x}}{2}$, $i(t) = \frac{2E}{L\sqrt{\Delta}} e^{-\frac{R}{2L}t} \text{sh}\left(\frac{\sqrt{\Delta}}{2}t\right)$. Si $\Delta < 0$, nous avons deux ra-

cines complexes conjuguées : $p_1 = \frac{-R}{2L} + i \frac{\sqrt{-\Delta}}{2}$ et $p_2 = \frac{-R}{2L} - i \frac{\sqrt{-\Delta}}{2}$ donc

$i(t) = \frac{E}{iL\sqrt{-\Delta}} \left(e^{-\frac{R}{2L}t} \left(e^{\frac{i\sqrt{-\Delta}}{2}t} - e^{-\frac{i\sqrt{-\Delta}}{2}t} \right) \right) = \frac{E}{L} e^{-\frac{R}{2L}t} \frac{2}{\sqrt{-\Delta}} \sin\left(\frac{\sqrt{-\Delta}}{2}t\right)$. Si $\Delta = 0$, nous

avons une racine réelle double : $I(p) = \frac{E}{L} \frac{1}{\left(p + \frac{R}{2L}\right)^2}$ car $p_1 = p_2 = \frac{-R}{2L}$ et ainsi :

$i(t) = \frac{E}{L} t e^{-\frac{R}{2L}t}$.

2. Nous avons $p^2X(p) - px_0 - v_0 + k^2X(p) = \frac{Aw}{p^2 + w^2}$ puis, $X(p) =$

$\frac{Aw}{k^2 - w^2} \left(\frac{1}{p^2 + w^2} - \frac{1}{p^2 + k^2} \right) + x_0 \frac{p}{p^2 + k^2} + \frac{v_0}{p^2 + k^2}$ et, par conséquent, nous obtenons :

$$x(t) = \frac{A}{k^2 - w^2} \sin(wt) + x_0 \cos(kt) + \left(\frac{v_0}{k} - \frac{Aw}{k(k^2 - w^2)} \right) \sin(kt).$$

Pour s'entraîner (solutions p. 540)

69.1 Calculer la transformée de Laplace inverse de $X(p) = \frac{p+3}{(p-2)(p+5)}$.

69.2 Calculer la transformée de Laplace inverse de $X(p) = \frac{p}{(p^2+1)^2}$.

69.3 Calculer la transformée de Laplace inverse de $X(p) = \frac{2p-1}{p(p^2+3)(p+5)}$.

69.4 Soit $F(p) = \ln\left(1 + \frac{1}{p^2}\right)$, calculer $F'(p)$ et en déduire l'original de $F(p)$.

Exemples d'équations aux dérivées partielles

1. Équations aux dérivées partielles

Définition 70.1

Une équation aux dérivées partielles (EDP) est une relation liant une fonction de plusieurs variables et ses dérivées partielles par rapport aux différentes variables. Il n'y a pas, malheureusement, de technique générale pour résoudre les équations de n'importe quel type. Une EDP est d'ordre k si elle ne fait intervenir que des dérivées partielles d'ordre inférieur ou égal à k .

2. EDP du premier ordre

Définition 70.2

Soit f une fonction de n variables $x_1, x_2, \dots, x_n$ continue et ayant des dérivées partielles sur un ouvert U de $\mathbb{R}^n$. Nous appelons équation aux dérivées partielles du premier ordre une relation de la forme :

$$F\left(x_1, x_2, \dots, x_n, f, \frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n}\right) = 0$$

entre les variables $x_1, x_2, \dots, x_n$, la fonction f et ses dérivées partielles du premier ordre et où F est une fonction connue.

3. EDP du deuxième ordre

Définition 70.3

Soit f une fonction de deux variables x et y continue et ayant des dérivées partielles sur un ouvert U de $\mathbb{R}^2$. Nous appelons équation aux dérivées partielles du deuxième ordre une relation de la forme :

$$F\left(x, y, f, \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial^2 f}{\partial x^2}, \frac{\partial^2 f}{\partial y^2}, \frac{\partial^2 f}{\partial x \partial y}\right) = 0$$

entre les variables x et y , la fonction f et ses dérivées partielles jusqu'à l'ordre deux et où F est une fonction connue.



Nous ferons l'étude de quelques EDP dans des cas bien particuliers se ramenant à la physique.

1. Température d'une plaque d'acier. Une grande plaque d'acier d'épaisseur $2L$, de température initiale élevée T_0 est brusquement plongée dans un bain d'huile froide de température T_b pour obtenir une augmentation de sa dureté. La température varie en suivant l'équation aux dérivées partielles :

$$\frac{\partial T}{\partial t} = a \frac{\partial^2 T}{\partial x^2}.$$

où ρ est la masse volumique de l'acier (constante), c_p la capacité thermique massique (constante), λ la conductivité thermique (constante), a la diffusivité de l'acier (constante) et $a = \frac{\lambda}{\rho c_p}$.

Résoudre cette EDP.

2. Élongation transversale d'une corde.

Soit une corde de longueur l , l'élongation transversale z d'un point M d'abscisse x est solution de l'EDP :

$$\frac{\partial^2 z}{\partial x^2} - \frac{1}{a^2} \frac{\partial^2 z}{\partial t^2} = 0, \quad z(0, t) = z(l, t) = 0.$$

Résoudre l'EDP.

1. Les opérateurs $\frac{\partial}{\partial t}$ et $\frac{\partial^2}{\partial x^2}$ sont séparés dans l'équation aux dérivées partielles, nous allons donc utiliser la méthode de séparation des variables. Nous posons : $T(x, t) = f(t)g(x)$ donc l'équation initiale devient $f'(t)g(x) = af(t)g''(x)$ d'où $\frac{f'(t)}{f(t)} = a \frac{g''(x)}{g(x)}$.

Le premier membre de l'équation est fonction de t , alors que le second membre est fonction de x . Donc, nous avons : $\frac{f'(t)}{f(t)} = k$ et $\frac{g''(x)}{g(x)} = \frac{k}{a}$, ce qui s'écrit sous la forme :

$$\begin{cases} f'(t) - kf(t) = 0 & (1) \\ g'(x) - \frac{k}{a}g(x) = 0. & (2) \end{cases}$$

Soit $A \in \mathbb{R}$, $f(t) = Ae^{kt}$ est solution de (1). Or il faut que la température tende vers 0 en $+\infty$ donc $k < 0$, nous posons alors $k = -\beta^2$.

Par conséquent (2) devient $\Leftrightarrow g''(x) + \frac{\beta^2}{a}g(x) = 0$. C'est une équation différentielle du second ordre homogène à coefficients constants. L'équation caractéristique s'écrit : $r^2 + \frac{\beta^2}{a} = 0$.

Ainsi, $g(x) = B \cos\left(\frac{\beta}{\sqrt{a}}x\right) + C \sin\left(\frac{\beta}{\sqrt{a}}x\right)$, où B et C sont des constantes réelles.

La solution générale est alors :

$$T(x, t) = \left(K_1 \cos\left(\frac{\beta}{\sqrt{a}}x\right) + K_2 \sin\left(\frac{\beta}{\sqrt{a}}x\right) \right) e^{-\beta^2 t}$$

avec $K_1 = AB$ et $K_2 = AC$.



Déterminons maintenant K_1 et K_2 .

Nous avons $T(0, t) = 0 \Leftrightarrow K_1 = 0$ et $T(2L, t) = 0 \Leftrightarrow K_2 \sin\left(\frac{\beta}{\sqrt{a}}2L\right) e^{-\beta^2 t} = 0$ donc il

existe $n \in \mathbb{Z}$ tel que $\frac{\beta}{\sqrt{a}}2L = n\pi \Leftrightarrow \beta = \frac{n\pi\sqrt{a}}{2L}$. Par conséquent, nous avons

$$T(x, t) = K_2 \sin\left(\frac{n\pi x}{2L}\right) e^{-\frac{n^2\pi^2 at}{4L^2}}.$$

Fiche 56



L'équation de la chaleur est linéaire, essayons d'écrire la température recherchée comme une combinaison linéaire infinie, c'est-à-dire une série, de toutes les solutions que nous venons de trouver. Nous avons donc : $T(x, t) = \sum_{n=1}^{+\infty} K_n \sin\left(\frac{n\pi x}{2L}\right) e^{-\frac{n^2 \pi^2 a t}{4L^2}}$.

Calculons K_n . Pour ce faire nous admettons qu'il est possible d'intervertir le symbole $\sum$ et l'intégrale.

Nous avons : $T(x, 0) = T_0 - T_b$ pour $x \in]0; 2L[$, et par conséquent $T(x, 0) \sin\left(\frac{m\pi x}{2L}\right) = (T_0 - T_b) \sin\left(\frac{m\pi x}{2L}\right)$, avec m un entier.

En intégrant, nous avons : $\int_0^{2L} (T_0 - T_b) \sin\left(\frac{m\pi x}{2L}\right) dx = \frac{2L(T_0 - T_b)}{m\pi} (1 - \cos(m\pi))$. Pour $n \neq m$:

$$\sum_{n=1}^{+\infty} K_n \int_0^{2L} \sin\left(\frac{n\pi x}{2L}\right) \sin\left(\frac{m\pi x}{2L}\right) dx = \sum_{n=1}^{+\infty} \frac{K_n}{2} \int_0^{2L} \left(\cos\left(\frac{(n-m)\pi x}{2L}\right) - \cos\left(\frac{(n+m)\pi x}{2L}\right) \right) dx = 0.$$

Pour $n = m$, nous avons : $\int_0^{2L} \sin^2\left(\frac{m\pi x}{2L}\right) dx = \int_0^{2L} \frac{1 - \cos\left(\frac{m\pi x}{L}\right)}{2} dx = L$.

Donc $K_m L = \frac{2L(T_0 - T_b)}{m\pi} (1 - \cos(m\pi))$.

Si $m = 2p$ est pair, $\cos(m\pi) = \cos(0) = 1$ donc $K_m L = 0$. Si $m = 2p + 1$ est impair, $\cos(m\pi) = \cos(\pi) = -1$ donc $K_m L = \frac{4L(T_0 - T_b)}{m\pi}$.

D'où $K_{2p+1} = \frac{4(T_0 - T_b)}{(2p+1)\pi}$, avec $p \in \mathbb{N}$.

$$T(x, t) = 4(T_0 - T_b) \sum_{p=0}^{+\infty} \frac{1}{(2p+1)\pi} \sin\left(\frac{(2p+1)\pi}{2L} x\right) e^{-\frac{(2p+1)^2 \pi^2 a t}{4L^2}}.$$

À l'instant initial ($t = 0$) et au milieu de la plaque ($x = L$), nous avons :

$$T(L, 0) = 4(T_0 - T_b) \sum_{p=0}^{+\infty} \frac{1}{(2p+1)\pi} \sin\left(\frac{(2p+1)\pi}{2}\right) = 4(T_0 - T_b) \sum_{p=0}^{+\infty} \frac{1}{(2p+1)\pi} (-1)^p.$$

Or, $\arctan(x) = \sum_{p=1}^{+\infty} \frac{(-1)^{p-1} x^{2p-1}}{2p-1} = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \dots$. Ici $x = 1$ et $\arctan(1) = \frac{\pi}{4}$. Ainsi, $T(L, 0) = T_0 - T_b$.

2. Nous cherchons une solution de la forme : $z(x, t) = g(x)f(t)$.

Nous avons alors $g''(x)f(t) - \frac{1}{a^2}g(x)f''(t) = 0 \iff \frac{g''(x)}{g(x)} = \frac{1}{a^2} \frac{f''(t)}{f(t)} = \lambda$ avec $\lambda < 0$. Ainsi

$$\begin{cases} g''(x) - \lambda g(x) = 0 & (1) \\ f''(t) - a^2 \lambda f(t) = 0. & (2) \end{cases}$$



Fiche 51

Par conséquent, $g(x) = A_1 \cos(\sqrt{-\lambda}x) + A_2 \sin(\sqrt{-\lambda}x)$ et $f(t) = B_1 \cos(a\sqrt{-\lambda}t) + B_2 \sin(a\sqrt{-\lambda}t)$.

$$z(x, t) = (A_1 \cos(\sqrt{-\lambda}x) + A_2 \sin(\sqrt{-\lambda}x))(B_1 \cos(a\sqrt{-\lambda}t) + B_2 \sin(a\sqrt{-\lambda}t)).$$

$z(0, t) = 0$ d'où $A_1 = 0$ et $z(l, t) = 0$ d'où $\sqrt{-\lambda}l = n\pi$, avec $n \in \mathbb{Z}$. Puis,

$$z(x, t) = \sin\left(\frac{n\pi x}{l}\right) \left(A \cos\left(\frac{n\pi at}{l}\right) + B \sin\left(\frac{n\pi at}{l}\right) \right)$$

avec $A_2 B_1 = A$ et $A_2 B_2 = B$.

Si la corde n'a pas de vitesse initiale : $\frac{\partial z}{\partial t}(x, 0) = 0 \iff B = 0$. Donc $z(t) = A \sin\left(\frac{n\pi x}{l}\right) \cos\left(\frac{n\pi at}{l}\right)$.



L'EDP est linéaire, nous pouvons donc prendre pour solution générale, sous réserve de convergence, une série formée des solutions précédentes.

$$z(x, t) = \sum_{n=1}^{+\infty} z_n(x, t) = \sum_{n=1}^{+\infty} A_n \sin\left(\frac{n\pi x}{l}\right) \cos\left(\frac{n\pi at}{l}\right).$$

Si $t = 0$, alors $z(x, 0) = \sum_{n=1}^{+\infty} A_n \sin\left(\frac{n\pi x}{l}\right) = f(x)$. Si, par exemple, cette série converge uniformément sur $[0; l]$, alors A_n est le coefficient de Fourier du développement de f en série de sinus : $A_n = \frac{2}{l} \int_0^l f(x) \sin\left(\frac{n\pi x}{l}\right) dx$.

$$z(x, t) = \sum_{n=1}^{+\infty} \left(\frac{2}{l} \int_0^l f(x) \sin\left(\frac{n\pi x}{l}\right) dx \right) \cos\left(\frac{n\pi at}{l}\right).$$

Pour s'entraîner (solutions p. 540)

70.1 En passant en coordonnées polaires, déterminer f telle que :

$$x \frac{\partial f}{\partial x}(x, y) + y \frac{\partial f}{\partial y}(x, y) = \sqrt{x^3 + y^3}.$$

70.2 

Résoudre, en utilisant coordonnées polaires, l'équation aux dérivées partielles :

$$x \frac{\partial f}{\partial x} + y \frac{\partial f}{\partial y} = \frac{y}{x}.$$

70.3 En effectuant le changement de variable $u = x^2 - y^2 - 2xy$ et $v = y$, déterminer

f de classe $\mathcal{C}^1$ de $\mathbb{R}^2$ à valeurs dans $\mathbb{R}$ telle que :

$$(x + y) \frac{\partial f}{\partial x} + (x - y) \frac{\partial f}{\partial y} = 0.$$

70.4 Soit f une fonction de classe $\mathcal{C}^2$ de $\mathbb{R}^2$ à valeurs dans $\mathbb{R}$ vérifiant :

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = 0.$$

Montrer que l'application

$$\begin{aligned} \phi : r \in \mathbb{R}_+ &\rightarrow \mathbb{R} \\ r &\mapsto \int_0^{2\pi} f(r \cos(\theta), r \sin(\theta)) d\theta \end{aligned}$$

est constante.

La marque déposée « Wi-Fi », adoptée en 1999, est un ensemble de protocoles de communication sans fil qui permet de créer de véritables réseaux locaux gérant beaucoup d'utilisateurs en offrant des débits de 3 à 4 Mbits/s (Mégabits par seconde), le bit étant un chiffre binaire (0 ou 1). Le terme Wi-Fi provient du terme Hi-Fi en suggérant la contraction de Wireless Fidelity.

Les types d'équipement

Il existe différents types d'équipement pour la mise en place d'un réseau sans fil Wi-Fi :

- les adaptateurs sans fils ou cartes d'accès (« *wireless adapters* » ou « *network interface controller* ») ; il s'agit d'une carte réseau à la norme 802.11 permettant à une machine de se connecter à un réseau sans fil. Les adaptateurs Wi-Fi sont disponibles dans de nombreux formats : carte PCI, carte PCMCIA, adaptateur USB, carte CompactFlash...
- les points d'accès (« *access point* »), parfois appelés bornes sans fils permettant de donner un accès au réseau filaire (auquel il est raccordé) aux différentes stations avoisinantes équipées de cartes Wi-Fi.

Le standard 802.11 définit deux modes opératoires :

- le mode infrastructure dans lequel les clients sans fils sont connectés à un point d'accès. Il s'agit généralement du mode par défaut des cartes 802.11b.
- le mode ad hoc dans lequel les clients sont connectés sans aucun point d'accès.

Le mode infrastructure

Le Wi-Fi fonctionnant en mode infrastructure utilise les points d'accès, gérant l'ensemble des communications dans une même zone ou cellule. Dans un réseau à infrastructure et à plusieurs cellules, les bornes sont connectées entre elles par une liaison ou par un réseau filaire ou encore hertzien. Les terminaux peuvent alors se déplacer au sein de la cellule et garder une liaison directe avec le point d'accès, ou changer de cellule, ce qui s'appelle le roaming ou encore l'itinérance. Ce mode garantit un passage obligé par le point d'accès, il est donc possible de vérifier qui accède au réseau.

L'étalement du spectre par code utilisé dans le Wi-Fi met en œuvre les traitements suivants :

1. Le signal binaire des données ayant un débit de base $D = 1$ MHz est multiplié par une séquence d'un signal pseudo-aléatoire de débit plus élevé (11 MHz). Le signal résultant module la porteuse de l'émetteur en modulation de phase. À la réception, l'onde porteuse est démodulée et le résultat mélangé à la même séquence pseudo-aléatoire afin de récupérer les données binaires.
2. La modulation de phase en quadrature ou **QPSK** (Quadrature Phase Shift Keying) est utilisée dès que le débit dépasse 1 Mbits/s.

La modulation par déplacement de phase à quatre états : MDP-4 ou QPSK

Les bits du train binaire (succession de bits aléatoires) sont groupés par deux pour former des mots de deux bits. La phase de la porteuse peut prendre quatre valeurs et chaque valeur est associée à un mot binaire de deux bits. Dans le cas d'une modulation de phase à quatre états, la durée d'un symbole, notée T_s , est égale à deux fois la durée d'un bit, notée T_b , ou encore $T_s = 2T_b$. Le signal modulé s'écrit : $s(t) = \sum_k \left(A \cos(w_p t + \varphi_k) h(t - kT_s) \right)$, où $w_p = 2\pi f_p$, f_p la fréquence de la porteuse et φ_k correspond aux quatre valeurs que peut prendre la phase sur un intervalle de temps de durée T_s , à savoir : $\left\{ \frac{\pi}{4}, \frac{3\pi}{4}, \frac{5\pi}{4}, \frac{7\pi}{4} \right\}$. Ces quatre états de phase ont été

définis de telle sorte qu'ils soient répartis uniformément et espacés de $\frac{\pi}{2}$. $h(t)$ est la fonction de mise en forme associée à la modulation, elle a une durée égale à la durée d'un symbole T_s .

Remarque : la modulation QSPK, utilisée pour le Wi-Fi, est aussi employée pour la télévision numérique par satellite.

Maintenant, introduisons la densité spectrale de puissance du signal, notée $G(f)$, qui représente la répartition de la puissance d'un signal suivant les fréquences et qui

se définit par : $G(f) = \frac{A^2 T_s}{2} \left(\frac{\sin(\pi(f - f_p)T_s)}{\pi(f - f_p)T_s} \right)^2$. Ainsi $G(f)$ nous montre qu'une

bande passante égale à $\frac{2}{T_s}$ est considérée comme suffisante pour transmettre les données par QPSK. En effet, plus de 90 % de l'information est comprise dans cette bande passante.

Le risque sanitaire

Le Wi-Fi apparaît au moment où se développent des interrogations quant à l'impact des radiofréquences sur la santé de l'homme. Des débats scientifiques se sont multipliés autour du téléphone mobile, et le débat s'est étendu à l'ensemble des technologies radio reposant sur les micro-ondes, notamment les technologies GSM, la 3G, ou encore la 3G+ et le Wi-Fi. Les ondes émises par les équipements Wi-Fi se diffusent dans l'ensemble de l'environnement, mais la fréquence relativement élevée de ces ondes (2,4 GHz) les fait mal traverser les murs. En outre, la puissance émise par les équipements Wi-Fi (~30 mW) est vingt fois moindre que celle émise par les téléphones mobiles (~600 mW). De plus, le téléphone est généralement tenu à proximité immédiate du cerveau, ce qui n'est pas le cas des équipements Wi-Fi (à l'exception des téléphones Wi-Fi), ou, à une dizaine de centimètres, la densité de puissance du signal est déjà fortement atténuée. Plusieurs organismes ont réalisé des études au sujet de l'effet sur la santé de le Wi-Fi, et ont, dans un premier temps, majoritairement conclu qu'il n'y avait aucune raison de craindre que le Wi-Fi soit dangereuse pour la santé dans le cadre d'une utilisation normale.

Les transmissions de données sans fil vont encore progresser en utilisant des bandes de fréquences térahertz (10^{12} Hz). Nous pourrons bientôt transmettre des informations avec des débits dix fois supérieurs aux débits actuels.

Probabilités

Partie

3



Introduction

De nombreux domaines scientifiques sont liés à l'aléatoire : la finance et l'assurance, la météorologie et la climatologie, la survie d'une espèce, les files d'attente et les télécommunications, etc.

En vue de comprendre et de maîtriser cette composante aléatoire, nous disposons d'un langage commun à toutes ces sciences. Ce langage est celui des probabilités.

Dans cette partie, nous présentons le vocabulaire et les résultats qui constituent le socle commun de tous les scientifiques travaillant dans des domaines où l'aléatoire s'invite : définitions d'un événement (fiche 72), d'une mesure de probabilité (fiches 72 et 73), d'une variable aléatoire et de la loi de probabilité (fiche 75), loi des grands nombres (fiche 92), et théorème de la limite centrale (fiche 93). De nombreux exemples d'applications montrent comment manipuler ces notions de probabilités.

En fin de partie, un focus fait le lien entre probabilités et jeux de hasard.

Les bonus web sur Dunod.com

 Retrouvez les corrigés détaillés d'une sélection d'exercices sur le site dunod.com

L'objet de cette Fiche est de dénombrer les dispositions possibles de n objets distincts, avec $n \in \mathbb{N}^*$. Nous ne précisons pas la nature de ces objets. Ce peut-être des lettres, des boules, des résultats d'expériences, etc.

1. Permutations

Imaginons que nous voulons **ordonner** n objets distincts, c'est-à-dire ranger ces objets en attribuant à chacun une position exclusive de 1 à n .

Définition 71.1

Le nombre de **permutations** est le nombre de manières d'ordonner n objets distincts. Nous notons P_n ce nombre et

$$P_n = n!$$



Point méthode

Ce résultat provient de la démarche suivante :

1. parmi les n objets, nous choisissons celui qui prendra la position 1 : il y a n choix possibles,
2. parmi les $n - 1$ objets restants, nous choisissons celui qui prendra la position 2 : il y a $n - 1$ choix possibles,
3. parmi les $n - 2$ objets restants, nous choisissons celui qui prendra la position 3 : il y a $n - 2$ choix possibles,
4. etc.
5. À la $n^{\text{ème}}$ itération, il ne reste qu'un objet pour prendre la position n : il y a un seul choix possible,

d'où

$$P_n = n \times (n - 1) \times \dots \times 1 = n!$$



Nous prenons la convention $0! = 1$.

2. Arrangements sans remise ou sans répétition

Soit un ensemble de n objets distincts. Imaginons que nous voulons ranger un sous-ensemble de $1 \leq k \leq n$ de ces objets en attribuant à chaque objet de ce sous-ensemble une position exclusive de 1 à k .

Définition 71.2

Soient deux entiers $n \in \mathbb{N}^*$ et $k \leq n$. Le nombre d'**arrangements sans remise** ou **sans répétition** de k objets pris parmi n objets distincts est le nombre de manières d'ordonner ces k objets. Nous notons A_n^k ce nombre et

$$A_n^k = \frac{n!}{(n - k)!}.$$



Point méthode

La formule donnant A_n^k provient de la démarche suivante :

1. parmi les n objets, nous choisissons celui qui prendra la position 1 : il y a n choix possibles,
2. parmi les $n - 1$ objets restants, nous choisissons celui qui prendra la position 2 : il y a $n - 1$ choix possibles,
3. parmi les $n - 2$ objets restants, nous choisissons celui qui prendra la position 3 : il y a $n - 2$ choix possibles,
4. etc.
5. Parmi les $n - (k - 1)$ objets restants, nous choisissons celui qui prendra la position k : il y a $n - k + 1$ choix possibles,

d'où

$$A_n^k = n \times (n - 1) \times \dots \times (n - (k - 1)) = \frac{n!}{(n - k)!}.$$



- Le nombre d'arrangements sans remise ou sans répétition est également appelé nombre d'**arrangements**.
- Si $k = n$, nous rangeons tout l'ensemble des n objets, donc $A_n^n = P_n = n!$
- Si $k = 0$, nous prenons la convention $A_n^0 = 1$.

3. Combinaisons

Soit un ensemble de n objets distincts. Imaginons que nous voulons extraire un sous-ensemble de $k \leq n$ de ces objets (sans affecter d'ordre à ces k objets).

Définition 71.3

Soient deux entiers $n \in \mathbb{N}^*$ et $k \leq n$. Le nombre de **combinaisons** de k objets pris parmi n objets distincts est le nombre d'ensembles de k objets pris parmi les n objets de départ. Nous notons $\binom{n}{k}$ ou C_n^k ce nombre et

$$\binom{n}{k} = \frac{n!}{k!(n - k)!} = \frac{A_n^k}{P_k}.$$



Point méthode

La formule donnant $\binom{n}{k}$ provient de la démarche suivante :

1. nous écrivons sur des bouts de papier tous les arrangements possibles de k objets pris parmi les n objets : il y en a A_n^k ,
2. nous regroupons par lots les bouts de papier arrangeant les k mêmes objets,
3. chaque lot est constitué de permutations de k objets : chaque lot est donc constitué de $k!$ bouts de papier,
4. le nombre de combinaisons $\binom{n}{k}$ est le nombre de lots obtenus : c'est-à-dire

$$\binom{n}{k} = \frac{A_n^k}{P_k}.$$

Théorème 71.1

Soient deux entiers $n \in \mathbb{N}^*$ et $1 \leq k \leq n$, nous avons la formule du **Triangle de Pascal** :

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}.$$

4. Arrangements avec remise ou avec répétitions

Soit un ensemble de $n \in \mathbb{N}^*$ objets distincts. Imaginons que nous prenons $k \in \mathbb{N}$ objets, **avec remise** dans cet ensemble, puis que nous **ordonnons** ces k objets.

Définition 71.4

Le nombre d'**arrangements avec remise** ou **avec répétitions** de $k \in \mathbb{N}$ objets pris avec remise parmi $n \in \mathbb{N}^*$ objets distincts est le nombre de manières d'ordonner ces k objets. Nous notons B_n^k ce nombre et

$$B_n^k = n^k.$$



Point méthode

La formule donnant B_n^k provient de la démarche suivante :

1. parmi les n objets, nous choisissons celui qui prendra la position 1 : il y a n choix possibles, puis nous le remettons avec ses $n - 1$ confrères,
2. parmi les n objets, nous choisissons celui qui prendra la position 2 : il y a n choix possibles, puis nous le remettons avec ses $n - 1$ confrères,
3. etc.
4. Parmi les n objets, nous choisissons celui qui prendra la position k : il y a n choix possibles, puis nous le remettons avec ses $n - 1$ confrères,

d'où

$$B_n^k = \underbrace{n \times n \times \dots \times n}_{k \text{ fois}} = n^k.$$

5. Ensembles dénombrables

Nous admettrons qu'un **ensemble dénombrable** est un ensemble tel que tous ses éléments peuvent être numérotés. Par exemple, $\mathbb{N}$ ou $\mathbb{Z}$ sont des ensembles dénombrables, mais $\mathbb{R}$ ne l'est pas.

Application

Dénombrement

Trois personnes lancent un dé chacune leur tour.

1. Combien y-a-t-il de résultats possibles ?
2. Combien y-a-t-il de cas où toutes les personnes ont des résultats différents ?
3. Combien y-a-t-il de cas où toutes les personnes ont le même résultat ?
4. Combien y-a-t-il de cas où deux personnes au moins ont le même résultat ?

1. Le dé a 6 faces. Les trois personnes lancent le dé, le nombre d_1 de résultats est un arrangement avec remise, ou plus précisément $d_1 = B_6^3 = 6^3 = 216$.
2. Le nombre d_2 de cas avec des résultats différents est un arrangement de 3 faces parmi 6 faces possibles du dé, c'est-à-dire $d_2 = A_6^3 = \frac{6!}{3!} = 6 \times 5 \times 4 = 120$.
3. Le nombre d_3 de cas avec des résultats identiques est une combinaison de 1 face parmi 6 faces possibles du dé, c'est-à-dire $d_3 = C_6^1 = \frac{6!}{1!5!} = 6$.
4. Le nombre c de cas où deux personnes au moins ont le même résultat est égal au nombre de cas où les trois personnes ont le même résultat, plus le nombre de cas où deux personnes ont le même résultat et la troisième a un résultat différent des deux autres.



Il est plus simple ici de dire que le nombre d_1 de résultats possibles est égal au nombre d_2 de cas avec des résultats tous différents plus le nombre de cas c où deux personnes au moins ont le même résultat. Nous en déduisons que $c = d_1 - d_2 = 216 - 120 = 96$.

Pour s'entraîner (solutions p. 541)

71.1 En France, chaque véhicule possède un numéro composé de sept caractères alphanumériques : deux lettres, un tiret, un nombre de trois chiffres (de 001 à 999), un tiret et deux lettres (exemple : CF-854-AB). Les lettres aujourd'hui interdites sont les I, O et U du fait de leur trop proche ressemblance avec le 1, le 0 et le V, ainsi que les séries SS et WW du bloc de gauche et la série SS du bloc de droite. Montrer que le nombre de plaques d'immatriculation autorisées par ce système est 277 977 744.

71.2 En Formule 1, 22 pilotes sont en course. Combien y-a-t-il de listes ordonnées des 10 premiers pilotes à l'arrivée ? Combien y-a-t-il de listes non ordonnées des 10 premiers pilotes à l'arrivée ?

71.3 1. Le bureau des élèves d'une école est constitué d'un président, d'un trésorier, et

d'un secrétaire. Combien de bureaux peuvent ainsi être formés si l'école a 500 élèves ?

2. Pour faire passer un oral à un groupe de 20 étudiants, le professeur tire l'ordre de passage des étudiants au hasard, et inscrit l'ordre sur une liste. Combien y-a-t-il de listes possibles ?

71.4 Trente demoiselles se présentent à l'élection de Miss France. Combien de classements Miss France, Première, Deuxième et Troisième Dauphines peuvent être formés ?

71.5 L'ADN est composé de bases azotées Adénine, Cytosine, Guanine et Thymine.

Un acide aminé est une séquence de trois bases. Quel est le nombre d'acides aminés possibles ?

Événements et probabilité

Cette Fiche présente les notions de base nécessaires pour définir la notion de « probabilité ».

Définition 72.1

- Une **expérience aléatoire** est une expérience, qui à chaque réalisation dans des conditions identiques, peut donner des résultats différents. Dans la suite de cette définition, nous considérons une expérience aléatoire donnée.
- L'**univers** est l'ensemble de tous les résultats possibles de l'expérience aléatoire. L'univers sera noté Ω . Cet ensemble peut être fini, dénombrable ou non dénombrable. Dans la suite de cette définition, nous considérons une expérience aléatoire donnée et son univers Ω .
- Un **événement** A est un sous-ensemble de résultats possibles de l'expérience aléatoire, c'est donc une partie de Ω et nous notons $A \subset \Omega$. L'ensemble des événements est noté $\mathcal{P}(\Omega)$: c'est aussi l'**ensemble des parties** de Ω .
- L'**événement impossible** ou **vide** ne contient aucun résultat possible de l'expérience aléatoire, il est noté $\emptyset$.
- L'**événement certain** contient tous les résultats possibles de l'expérience aléatoire. C'est Ω .
- Un **événement élémentaire** ou **éventualité** est un singleton de Ω , c'est-à-dire un événement qui contient un seul résultat possible de l'expérience aléatoire. Un événement élémentaire est noté $\{\omega\}$, où $\omega \in \Omega$.

Les événements étant des ensembles, nous avons les définitions suivantes :

 Fiche 3

Définition 72.2

Soient Ω un univers, et A, B deux événements de Ω .

- L'**intersection** des événements A et B est l'événement $A \cap B = \{\omega \in \Omega; \omega \in A \text{ et } \omega \in B\}$.
- L'**union** des événements A et B est l'événement $A \cup B = \{\omega \in \Omega; \omega \in A \text{ ou } \omega \in B\}$.
- La **soustraction** de A à B est l'événement $B \setminus A = \{\omega \in \Omega; \omega \in B \text{ et } \omega \notin A\}$.
- L'**événement complémentaire** de A dans Ω est l'événement $\bar{A} = \Omega \setminus A = \{\omega \in \Omega; \omega \notin A\}$. Nous avons $A \cup \bar{A} = \Omega$. L'événement complémentaire est parfois noté A^c .
- L'**inclusion** de l'événement A dans l'événement B , notée $A \subset B$, signifie : $\omega \in A \implies \omega \in B$. Nous disons que l'événement A est **inclus** dans l'événement B .
- Deux événements A et B de l'univers Ω sont **disjoints** ou **incompatibles** si $A \cap B = \emptyset$.



Tout événement A et son complémentaire $\bar{A}$ sont disjoints.

Définition 72.3

Soient Ω un univers, et $(A_i)_{i \in I}$ une famille dénombrable d'événements de Ω .

- L'**intersection** des événements $(A_i)_{i \in I}$ est l'événement $\bigcap_{i \in I} A_i = \{\omega \in \Omega; \forall i \in I, \omega \in A_i\}$.
- L'**union** des événements $(A_i)_{i \in I}$ est l'événement $\bigcup_{i \in I} A_i = \{\omega \in \Omega; \exists i \in I, \text{ tel que } \omega \in A_i\}$.

Exemple

Voici trois **expériences aléatoires** :

1. Lancer d'un dé à six faces non truqué.
2. Comptage du nombre de clients au guichet d'une banque, entre 10 h et 12 h.
3. Durée de vie d'une ampoule.

Les **univers** correspondants sont respectivement $\Omega_1 = \{1; 2; 3; 4; 5; 6\}$, $\Omega_2 = \mathbb{N}$, et $\Omega_3 = \mathbb{R}^+$.

Pour les trois expériences aléatoires précédentes, donnons trois **événements**. L'événement « la face du dé est impaire » correspond à la partie $A_1 = \{1; 3; 5\}$. L'événement « moins de 7 clients au guichet » correspond à la partie $A_2 = \{0; 1; 2; 3; 4; 5; 6; 7\}$. L'événement « l'ampoule vit plus de 50,2 jours » correspond à la partie $A_3 = [50, 2; +\infty[$.

Donnons maintenant les **événements complémentaires** des événements précédents. L'événement complémentaire de A_1 est « la face du dé est paire », il correspond à $\bar{A}_1 = \{2; 4; 6\}$. L'événement complémentaire de A_2 est « plus de 8 clients au guichet », il correspond à $\bar{A}_2 = \{8; 9; 10; \dots\}$. L'événement complémentaire de A_3 est « l'ampoule vit strictement moins de 50,2 jours », il correspond à $\bar{A}_3 = [0; 50, 2[$.

Définition 72.4

Soit $\mathcal{P}(\Omega)$ l'ensemble des parties de Ω . Une **mesure de probabilité** ou encore **probabilité** $\mathbb{P}$ est une application de $\mathcal{P}(\Omega)$ dans $[0; 1]$:

$$\mathbb{P} : \begin{cases} \mathcal{P}(\Omega) \longrightarrow [0; 1] \\ A \longmapsto \mathbb{P}(A) \end{cases}$$

qui vérifie :

1. $\mathbb{P}(\emptyset) = 0$ et $\mathbb{P}(\Omega) = 1$,
2. pour toute suite $(A_n)_{n \in \mathbb{N}}$ d'événements deux à deux disjoints (c'est-à-dire $A_n \cap A_m = \emptyset$ si $n \neq m$), alors

$$\mathbb{P}\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \sum_{n=0}^{+\infty} \mathbb{P}(A_n).$$

Pour tout $A \in \mathcal{P}(\Omega)$, le nombre réel $\mathbb{P}(A)$ s'appelle **probabilité de l'événement A** .

Le triplet $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ est alors appelé un **espace de probabilité**.



Pour définir une probabilité, il est donc nécessaire de définir l'univers Ω et l'ensemble des événements $\mathcal{P}(\Omega)$. Dans le cas d'univers continus, la définition de l'ensemble de départ de $\mathbb{P}$ est en réalité plus compliquée que $\mathcal{P}(\Omega)$: nous nous contenterons dans ce livre de définir la mesure de probabilité $\mathbb{P}$ sur des intervalles ou des réunions finies ou dénombrables d'intervalles ouverts ou fermés de $\mathbb{R}$.

Théorème 72.1

- La probabilité de l'événement complémentaire d'un événement A est égale à $\mathbb{P}(\bar{A}) = 1 - \mathbb{P}(A)$.
- Si A et B sont deux événements alors $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.
- Si A et B sont deux événements disjoints alors $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$.
- Si $A_1, \dots, A_n$ sont des événements, alors $\mathbb{P}\left(\bigcup_{i=1}^n A_i\right) \leq \sum_{i=1}^n \mathbb{P}(A_i)$.
- Si $A_1, \dots, A_n$ sont des événements **disjoints deux à deux**, c'est-à-dire $A_i \cap A_j = \emptyset$ si $i \neq j$, alors $\mathbb{P}\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n \mathbb{P}(A_i)$.
- Si un événement A implique l'événement B , sa probabilité est plus petite : $A \subset B \implies \mathbb{P}(A) \leq \mathbb{P}(B)$.
- Si A et B sont deux événements alors $\mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A \cap B)$.

Définition 72.5

Soit I une partie de $\mathbb{N}$. La famille $(A_i)_{i \in I}$ d'événements de l'univers Ω forme un **système complet d'événements** ou une **partition de Ω** si :

1. $\bigcup_{i \in I} A_i = \Omega$,
2. pour $(i, j) \in I^2$ avec $i \neq j$, nous avons $A_i \cap A_j = \emptyset$.

Théorème 72.2

Soit $(A_i)_{i \in I}$ une partition de Ω . Alors pour tout événement B de Ω , nous avons :

$$\mathbb{P}(B) = \sum_{i \in I} \mathbb{P}(A_i \cap B).$$

Application **Un calcul de probabilité**

Dans un atelier, des pièces proviennent de deux machines a et b . Certaines sont défectueuses. Nous savons que 0,06 des pièces sont en provenance de la machine a et défectueuses. Nous savons également que 0,29 des pièces proviennent de la machine b et sont défectueuses.

Quelle est la probabilité qu'une pièce produite par l'atelier soit défectueuse ?

Modélisons le problème :

1. L'expérience aléatoire considérée est la provenance d'une pièce prise au hasard dans l'atelier (machine a ou b) et sa qualité (défectueuse ou non, ce que nous notons d ou $\bar{d}$).

2. Nous considérons l'**univers** $\Omega = \{a; b\} \times \{d; \bar{d}\}$. Toutes les pièces produites par l'atelier sont caractérisées par un couple $(\omega_1, \omega_2) \in \Omega$. Par exemple le couple (a, d) signifie que la pièce provient de la machine a et est défectueuse.

3. Écrivons les **événements** qui nous intéressent : soient A , B et D les événements $A = \ll \text{La pièce provient de la machine } a \gg$, $B = \ll \text{La pièce provient de la machine } b \gg$, $D = \ll \text{La pièce est défectueuse} \gg$. Par exemple, nous avons $A = \{(a, d), (a, \bar{d})\}$, ou encore, $A \cap D = \{(a, d)\}$.

4. Les événements A et B forment une **partition** de Ω : $A \cup B = \Omega$ et $A \cap B = \emptyset$. Remarquons au passage que $B = \bar{A}$. Le Théorème 72.2 donne donc

$$\mathbb{P}(D) = \mathbb{P}(D \cap A) + \mathbb{P}(D \cap B).$$

5. Traduisons les données de l'énoncé : nous avons $\mathbb{P}(A \cap D) = 0,06$ et $\mathbb{P}(B \cap D) = 0,29$. Finalement,

$$\mathbb{P}(D) = 0,06 + 0,29 = 0,35.$$

Pour s'entraîner (solutions p. 541)

72.1 Montrer que si $A_1, \dots, A_n$ est une partition de Ω , alors $\mathbb{P}(A_1 \cup \dots \cup A_n) = 1$.

72.2 Soient les événements $A = \ll \text{Demain il pleuvra} \gg$ et $B = \ll \text{Demain le soleil brillera toute la journée} \gg$, tels que $\mathbb{P}(A) = 0,2$ et $\mathbb{P}(B) = 0,5$. Pourquoi avons-nous $\mathbb{P}(A \cup B) = 0,7$?

72.3 Soient deux événements A et B tels que $\mathbb{P}(A) = 0,6$ et $\mathbb{P}(B) = 0,7$. Donner un encadrement de $\mathbb{P}(A \cap B)$.

72.4 M. Untel va voter lors d'un référendum. Les probabilités qu'il vote « oui » et qu'il vote « non » sont respectivement de 0,5 et 0,1. Quelle est la probabilité qu'il vote « blanc » ?

72.5 Décrire l'univers Ω associé à l'expérience aléatoire qui consiste à jeter deux dés. Quel est le cardinal de Ω ?

72.6 Soit une urne contenant 10 boules. Décrire l'univers Ω associé à l'expérience

aléatoire qui consiste à tirer deux boules l'une après l'autre, avec remise. Même question si les deux boules sont tirées simultanément.

72.7 Considérons une boîte d'allumettes, et les événements $A = \ll \text{la boîte contient entre 100 et 200 allumettes} \gg$ et $B = \ll \text{la boîte contient plus de 50 allumettes} \gg$. Montrer que $\mathbb{P}(A) \leq \mathbb{P}(B)$.

72.8 À l'arrivée d'un véhicule à un péage, considérons les événements $A_i = \ll \text{le véhicule contient } i \text{ personne(s)} \gg$, pour $i = 1; 2; 3$, $A_4 = \ll \text{le véhicule contient plus de 4 personnes} \gg$, $B = \ll \text{le véhicule est une moto} \gg$ et $C = \ll \text{le véhicule est un bus} \gg$. Nous donnons $\mathbb{P}(C \cap A_i) = 0,10$ pour $i = 1; 2; 3$, $\mathbb{P}(C \cap A_4) = 0,20$, et $\mathbb{P}(A_i) = 0,25$ pour $i = 1; 2; 3; 4$. Calculer $\mathbb{P}(C)$, $\mathbb{P}(\bar{C})$ et $\mathbb{P}(A_2 \cup A_3 \cup A_4)$. Pourquoi avons-nous $\mathbb{P}(B) \leq \mathbb{P}(\bar{C})$?

72.9 Soient A et B deux événements. Montrer que $\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B) = \mathbb{P}(\bar{A} \cap \bar{B}) - \mathbb{P}(\bar{A})\mathbb{P}(\bar{B})$.

Probabilité sur un univers fini, dénombrable ou continu

La définition de « probabilité » est une définition générale. Nous la détaillons ici dans le cas d'univers finis, dénombrables ou continus.

 Fiche 72

1. Probabilité sur un univers fini

Définition 73.1

Soient $N \in \mathbb{N}^*$ et Ω un **univers fini** de cardinal $\text{Card } \Omega = N$. Nous notons $\Omega = \{\omega_1; \dots; \omega_N\}$. Une probabilité $\mathbb{P}$ sur Ω est définie par la famille $(p_i)_{i=1, \dots, N}$ de réels compris entre 0 et 1 vérifiant $\sum_{i=1}^N p_i = 1$, et tels que pour tout $\omega_i \in \Omega$, la probabilité du singleton $\{\omega_i\}$ est donnée par $\mathbb{P}(\{\omega_i\}) = p_i$. La probabilité de tout événement $A \subset \Omega$ est alors définie par $\mathbb{P}(A) = \sum_{\omega_i \in A} \mathbb{P}(\{\omega_i\}) = \sum_{\omega_i \in A} p_i$.

Exemple

Soient 12 forages, chacun ayant une probabilité $p \in [0; 1]$ de contenir du pétrole. L'expérience aléatoire est le nombre de forages qui contiennent réellement du pétrole : $\Omega = \{0; 1; \dots; 12\}$. Pour tout $i \in \Omega$, la probabilité que i forages exactement contiennent du pétrole est $p_i = \binom{12}{i} p^i (1-p)^{12-i}$, où le coefficient $\binom{12}{i}$ a été défini à la Fiche 71.

Si Ω est un univers fini, nous avons le cas particulier suivant :

Définition 73.2

Soit Ω un univers fini, de cardinal noté $\text{Card } \Omega$. La **probabilité uniforme** ou **équiprobabilité** $\mathbb{P}$ sur Ω est la probabilité telle que tout événement élémentaire $\{\omega_i\}$ de Ω a la même probabilité :

$$\forall i \in \{1; \dots; \text{Card } \Omega\}, \quad p_i = \frac{1}{\text{Card } \Omega}.$$

Nous avons donc, pour tout événement A de Ω :

$$\mathbb{P}(A) = \sum_{\omega_i \in A} p_i = \frac{\text{Card } A}{\text{Card } \Omega}.$$

Dans le cas de l'**équiprobabilité sur un univers fini**, nous trouvons la formule

$$\mathbb{P}(A) = \frac{\text{nombre de résultats vérifiant } A}{\text{nombre total de résultats}}.$$


Exemple

Soit l'expérience aléatoire consistant à noter le numéro de la face obtenue lors d'un lancer de dé non truqué. Nous avons $\Omega = \{1; 2; 3; 4; 5; 6\}$ et, pour tout $i \in \Omega$, $\mathbb{P}(\{i\}) = 1/6 = 1/\text{Card } \Omega$. Chaque face ayant la même probabilité d'apparaître, $\mathbb{P}$ est la probabilité uniforme sur Ω .

2. Probabilité sur un univers dénombrable

La Définition 73.1 s'étend facilement au cas d'un univers dénombrable.

 Fiche 71

Définition 73.3

Une **probabilité sur un univers dénombrable** $\Omega = \{\omega_0; \omega_1; \dots; \omega_n; \dots\}$ est définie par la famille $(p_i)_{i \geq 0}$ de réels compris entre 0 et 1 vérifiant $\sum_{i \geq 0} p_i = 1$, et tels que la probabilité de tout singleton $\{\omega_i\}$ pour $i \in \mathbb{N}$ est donnée par $\mathbb{P}(\{\omega_i\}) = p_i$.

La probabilité de tout événement $A \subset \Omega$ est alors définie par $\mathbb{P}(A) = \sum_{\omega_i \in A} \mathbb{P}(\{\omega_i\}) = \sum_{\omega_i \in A} p_i$.

Exemple

Soit l'expérience aléatoire consistant à compter le nombre de désintégrations radioactives en une seconde en un lieu donné près d'une centrale nucléaire, le nombre moyen de désintégrations par seconde étant noté λ . Cette expérience est décrite par l'univers $\Omega = \mathbb{N}$, et, pour tout $i \in \mathbb{N}$, $\mathbb{P}(\{i\}) = e^{-\lambda} \frac{\lambda^i}{i!}$.

3. Densité de probabilité – Probabilité sur un univers continu

Soit un univers Ω continu : un intervalle de $\mathbb{R}$ ou $\mathbb{R}$ tout entier, un pavé de $\mathbb{R}^2$ ou $\mathbb{R}^2$ tout entier, un pavé de $\mathbb{R}^3$ ou $\mathbb{R}^3$ tout entier.

Définition 73.4

Soit Ω un univers continu comme ceux définis ci-dessus. Une **densité de probabilité** sur Ω est une fonction f définie sur Ω et à valeurs dans $\mathbb{R}$ telle que :

1. $\forall x \in \Omega, \quad f(x) \geq 0$,
2. $\int_{\Omega} f(x) dx = 1$.

La probabilité $\mathbb{P}$ sur Ω est définie par la formule $\mathbb{P}(A) = \int_A f(x) dx$, pour tout événement A de Ω . Cette probabilité est appelée **probabilité de densité** f .



- Par la suite, nous utiliserons le terme de densité pour parler de densité de probabilité.
- Dans le cas d'un univers fini ou dénombrable, la probabilité est définie via des sommes de réels p_i . Dans le cas d'un univers continu, la probabilité est souvent définie via des intégrales d'une densité. Dans ce livre, exceptés l'Application de la Fiche 80 et l'Exercice 80.6, nous ne considérerons que ce cas.
- Dans le cas d'un univers continu, si la probabilité est définie via une densité, la probabilité de tout événement élémentaire $\{\omega\} \subset \Omega$ est nulle car $\mathbb{P}(\{\omega\}) = \int_{\{\omega\}} f(x) dx = 0$.

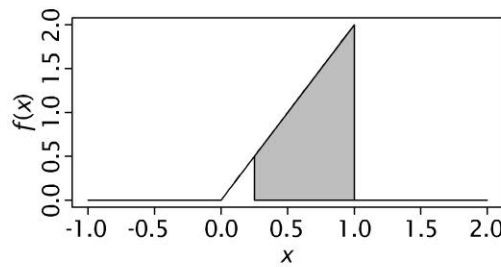


Figure 73.1 – Densité de probabilité $f(x) = 2x$ sur $\Omega = [0; 1]$. L'aire sous la courbe est égale à 1. La densité de probabilité est une fonction positive sur son domaine de définition. L'aire grisée est égale à $\mathbb{P}([0, 25; 1]) = \mathbb{P}([0, 25; +\infty]) = \mathbb{P}([0, 25; 1]) = \mathbb{P}([0, 25; +\infty])$.

Si Ω est un **intervalle de $\mathbb{R}$** , nous avons le cas particulier suivant :

Définition 73.5

Soit Ω un intervalle de $\mathbb{R}$, de longueur $\ell(\Omega)$ finie. La **probabilité uniforme** $\mathbb{P}$ sur Ω est la probabilité de densité $f(x) = \frac{\mathbb{1}_{x \in \Omega}}{\ell(\Omega)}$ où $\mathbb{1}_{x \in \Omega} = 1$ si $x \in \Omega$ et 0 si $x \notin \Omega$. Ainsi, pour tout événement A inclus dans Ω , nous avons :

$$\mathbb{P}(A) = \int_A f(x)dx = \int_A \frac{\mathbb{1}_{x \in \Omega}}{\ell(\Omega)} dx = \frac{\ell(A)}{\ell(\Omega)},$$

où $\ell(A)$ désigne la longueur de l'événement A .



Cette définition ressemble à la Définition 73.2. Nous remplaçons le cardinal par la longueur. De même, nous pouvons remplacer Ω intervalle de $\mathbb{R}$ par Ω un pavé de $\mathbb{R}^2$ ou $\mathbb{R}^3$. Dans ce cas, $\ell(\Omega)$ désigne respectivement l'aire et le volume de Ω .

Application Un calcul de probabilité

Nous lançons un dé non biaisé jusqu'à l'apparition de la face « 1 » et nous nous intéressons au nombre de lancers à effectuer avant l'apparition du « 1 ». Quelle est la probabilité que la face « 1 » n'apparaisse pas pendant les dix premiers lancers ?

L'**expérience aléatoire** est le comptage du nombre de lancers du dé jusqu'à l'apparition du « 1 ». L'**univers** Ω est $\mathbb{N}^*$. C'est un ensemble dénombrable. Le résultat $\omega_i = i$ est le résultat où la face « 1 » apparaît pour la première fois au $i^{\text{ème}}$ lancer. Décrivons la **probabilité** $\mathbb{P}$ associée à cette expérience aléatoire. Nous notons $p_i = \mathbb{P}(\{i\})$. Puisque le dé est non biaisé, et que la face « 1 » n'apparaît pas au cours des $i - 1$ premiers lancers, et seulement au $i^{\text{ème}}$ lancer, nous avons $p_i = \left(\frac{5}{6}\right)^{i-1} \times \frac{1}{6} = \frac{5^{i-1}}{6^i}$. Écrivons maintenant l'**événement** $B = \text{« ne pas tomber sur « 1 » pendant les 10 premiers lancers »}$, c'est-à-dire $B = \{11; 12; \dots\}$. Nous avons alors $\mathbb{P}(B) = \sum_{i \geq 11} p_i = \sum_{i \geq 11} \left(\frac{5}{6}\right)^{i-1} \times \frac{1}{6} = \frac{1}{6} \times \left(\frac{5}{6}\right)^{10} \times \sum_{i \geq 0} \left(\frac{5}{6}\right)^i = \frac{5^{10}}{6^{11}} \times \frac{1}{1 - \frac{5}{6}} = \left(\frac{5}{6}\right)^{10} \simeq 0,16$.



Pour le calcul, il est plus facile de calculer la probabilité de l'**événement complémentaire** $\bar{B}$ puis de calculer $\mathbb{P}(B) = 1 - \mathbb{P}(\bar{B})$.



Fiche 46

Ici, $\bar{B} = \{1; \dots; 10\}$, et $\mathbb{P}(\bar{B}) = \sum_{i=1}^{10} p_i$, et finalement $\mathbb{P}(B) = 1 - \sum_{i=1}^{10} p_i$.



Une valeur numérique approchée peut-être obtenue avec un logiciel de calcul (par exemple le langage **R**), nous trouvons $\mathbb{P}(B) \approx 0,16$.

Pour s'entraîner (solutions p. 542)

73.1 Soient un dé biaisé, et ci-dessous l'univers et la probabilité correspondant au résultat d'un lancer du dé :

Ω	1	2	3	4	5	6
p_i	p	$\frac{1}{6}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{4}$	$\frac{1}{3}$

Déterminer le réel p , puis la probabilité d'obtenir un numéro pair au lancer du dé.

73.2 Le loto consiste à tirer au sort 6 numéros distincts sur une grille de 49 numéros possibles. Le gagnant est celui qui donne les 6 bons numéros (peu importe l'ordre). J'achète une grille de loto. Quelle est la probabilité que je gagne ?

73.3 Dans une classe de n élèves, quelle est la probabilité qu'au moins deux élèves soient nés le même jour ? Nous négligeons les années bissextiles : nous supposons que l'année est toujours constituée de 365 jours.

73.4 Soit la famille de réels $(p_k)_{k \geq 1}$ tels que pour tout $k \geq 1$, $p_k = \frac{\alpha}{k!}$, avec α une

constante réelle positive fixée. Que vaut la constante α pour que cette famille définisse une probabilité sur $\mathbb{N}^*$?

73.5 Sur $\Omega = \mathbb{R}_+$, nous définissons la fonction $f(x) = Cxe^{-x^2} \mathbb{1}_{x \in \Omega}$ où nous rappelons que $\mathbb{1}_{x \in \Omega} = 1$ si $x \in \Omega$ et 0 sinon. Déterminer la constante C pour que la fonction f soit une densité de probabilité. Quelle est l'allure de f ? Calculer la probabilité de l'intervalle $[-2; 2]$. Même question pour les intervalles $[2; +\infty[$, $]-2; 2]$, $]-2; 2[$, et $]2; +\infty[$.

73.6 Une fléchette est jetée au hasard dans une cible circulaire de rayon 2 m. Quelle est la probabilité qu'elle tombe à moins de 1 m du centre de la cible ?

73.7 Un point est jeté au hasard dans le carré Ω de côté 1. Quelle est la probabilité que ce point tombe dans l'ensemble

$$B = \{(x, y) \in \Omega; x^2 + y^2 \geq 1 \text{ et } x \geq y\}?$$

Événements indépendants

Probabilité conditionnelle

Théorème de Bayes

Nous continuons ici d'introduire les bases fondamentales à la compréhension des probabilités. Nous définissons la probabilité conditionnelle d'un événement sachant un deuxième événement. Nous traitons le cas particulier, mais très important, où les deux événements sont indépendants. Enfin, le théorème de Bayes nous donnera une formule très utile pour calculer les probabilités conditionnelles.

1. Événements indépendants

Définition 74.1

Un événement A est **indépendant** d'un événement B si la réalisation de l'un n'est pas influencée par la réalisation de l'autre, ce qui se formule ainsi :

$$A \text{ et } B \text{ sont indépendants} \iff \mathbb{P}(A \cap B) = \mathbb{P}(A) \times \mathbb{P}(B).$$

La définition se généralise à n événements. Par exemple, A, B, C sont **mutuellement indépendants** si

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \times \mathbb{P}(B), \mathbb{P}(A \cap C) = \mathbb{P}(A) \times \mathbb{P}(C), \mathbb{P}(B \cap C) = \mathbb{P}(B) \times \mathbb{P}(C) \text{ et } \mathbb{P}(A \cap B \cap C) = \mathbb{P}(A) \times \mathbb{P}(B) \times \mathbb{P}(C).$$



Ne pas confondre **indépendant** et **disjoint** ou **incompatible** :

- A et B sont deux événements **indépendants** si $\mathbb{P}(A \cap B) = \mathbb{P}(A) \times \mathbb{P}(B)$,
- A et B sont deux événements **disjoints** ou **incompatibles** si $A \cap B = \emptyset$ et donc $\mathbb{P}(A \cap B) = 0$.

Théorème 74.1

Les événements A et B sont indépendants si et seulement si A et $\bar{B}$ le sont, ou si et seulement si $\bar{A}$ et $\bar{B}$ le sont.

2. Probabilité conditionnelle

Définition 74.2

Soit B un événement tel que $\mathbb{P}(B) \neq 0$. Pour tout événement A , la **probabilité conditionnelle** de A sachant que B est réalisé est définie par

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$



- La formule est souvent utilisée dans ce sens : $\mathbb{P}(A \cap B) = \mathbb{P}(B) \times \mathbb{P}(A|B)$.
- La formule de la Définition 74.1 n'est qu'un cas particulier correspondant au cas où A et B sont indépendants, c'est-à-dire au cas où $\mathbb{P}(A|B) = \mathbb{P}(A)$ (qui équivaut encore à $\mathbb{P}(B|A) = \mathbb{P}(B)$ si $\mathbb{P}(A) > 0$).

Exemple

Tirons au hasard une carte dans un jeu de 32 cartes. Sachant que la carte tirée est un cœur, la probabilité que la carte soit une dame est égale à $\frac{1}{8}$. En effet, l'univers Ω associé à l'expérience aléatoire est ici $\Omega = \ll \text{les 32 cartes du jeu} \gg$. Ω est muni de la probabilité $\mathbb{P}$ uniforme car le tirage est réalisé au hasard. Notons A l'événement « tirer une dame dans un jeu de 32 cartes », et B l'événement « tirer un cœur dans un jeu de 32 cartes ». L'événement $A \cap B$ est « tirer la dame de cœur dans un jeu de 32 cartes ». Nous avons $\mathbb{P}(A \cap B) = \frac{1}{32}$ et $\mathbb{P}(B) = \frac{1}{4}$, d'où $\mathbb{P}(A|B) = \frac{1}{8}$.

3. Théorème de Bayes

Théorème 74.2 Formule d'inversion

Soient A et B deux événements tels que $\mathbb{P}(A) > 0$ et $\mathbb{P}(B) > 0$, alors

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A) \times \mathbb{P}(B|A)}{\mathbb{P}(B)}.$$



- Cette formule d'inversion est très pratique lorsqu'il est plus facile de calculer $\mathbb{P}(B|A)$ que de calculer $\mathbb{P}(A|B)$.
- Ce résultat se retrouve très facilement : $\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(A) \times \mathbb{P}(B|A)}{\mathbb{P}(B)}$.
- Il peut arriver que le dénominateur $\mathbb{P}(B)$ ne soit pas connu directement. Il est possible de le calculer grâce à la formule des probabilités totales ci-dessous.

Théorème 74.3

Soient $I \subset \mathbb{N}$ et $(A_i)_{i \in I}$ une partition de Ω telle que $\mathbb{P}(A_i) > 0$, $\forall i \in I$. Alors pour tout événement B , nous avons la **formule des probabilités totales** :

$$\mathbb{P}(B) = \sum_{i \in I} \mathbb{P}(B|A_i) \times \mathbb{P}(A_i).$$

La formule d'inversion combinée avec la formule des probabilités totales donne le **Théorème de Bayes** : si $\mathbb{P}(B) > 0$, nous avons pour tout $k \in I$,

$$\mathbb{P}(A_k|B) = \frac{\mathbb{P}(B|A_k) \times \mathbb{P}(A_k)}{\sum_{i \in I} \mathbb{P}(B|A_i) \times \mathbb{P}(A_i)}.$$



Point méthode

Considérons la partition très simple donnée par A et $\bar{A}$ où A est un événement de Ω et $\bar{A}$ son complémentaire. La **formule des probabilités totales** s'écrit :

$$\mathbb{P}(B) = \mathbb{P}(B|A) \times \mathbb{P}(A) + \mathbb{P}(B|\bar{A}) \times \mathbb{P}(\bar{A}),$$

et le **Théorème de Bayes** donne :

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A) \times \mathbb{P}(A)}{\mathbb{P}(B|A) \times \mathbb{P}(A) + \mathbb{P}(B|\bar{A}) \times \mathbb{P}(\bar{A})}.$$

En 2009 en France, le taux de chômage parmi les femmes actives était de 0,094 et il était de 0,089 parmi les hommes actifs. Nous admettons que le taux d'hommes actifs est égal à 1,1 fois celui des femmes actives.

1. Calculer le taux de chômage.

2. Parmi les chômeurs, les hommes sont-ils plus nombreux que les femmes ?

L'univers Ω est ici l'ensemble de la population active française en 2009. Une **partition** de Ω est donnée par les événements F = « être une femme active », et $H = \bar{F}$ = « être un homme actif ». Considérons les événements C = « être au chômage » et $\bar{C}$ = « ne pas être au chômage ». Rappelons qu'être actif signifie être en âge de travailler (donc il est possible d'être actif et au chômage).

1. Le taux de chômage est $\mathbb{P}(C)$. D'après la **formule des probabilités totales** appliquée à la partition (F, H) de Ω , nous avons

$$\mathbb{P}(C) = \mathbb{P}(C|H) \times \mathbb{P}(H) + \mathbb{P}(C|F) \times \mathbb{P}(F).$$

L'énoncé nous indique que $\mathbb{P}(H) = 1,1 \times \mathbb{P}(F)$. Nous savons également que $\mathbb{P}(H) + \mathbb{P}(F) = 1$. Nous en déduisons que $\mathbb{P}(F) \times (1,1 + 1) = 1$, c'est-à-dire $\mathbb{P}(F) = \frac{1}{2,1} = 0,476$.

De plus, l'énoncé nous indique que $\mathbb{P}(C|F) = 0,094$ et $\mathbb{P}(C|H) = 0,089$. Nous concluons donc que le taux de chômage est

$$\mathbb{P}(C) = 0,089 \times 1,1 \times 0,476 + 0,094 \times 0,476 = 0,091.$$

2. D'après le **Théorème de Bayes**, nous avons :

$$\mathbb{P}(F|C) = \frac{\mathbb{P}(C|F) \times \mathbb{P}(F)}{\mathbb{P}(C|F) \times \mathbb{P}(F) + \mathbb{P}(C|H) \times \mathbb{P}(H)}$$

et

$$\mathbb{P}(H|C) = \frac{\mathbb{P}(C|H) \times \mathbb{P}(H)}{\mathbb{P}(C|F) \times \mathbb{P}(F) + \mathbb{P}(C|H) \times \mathbb{P}(H)}.$$

Nous avons donc

$$\frac{\mathbb{P}(H|C)}{\mathbb{P}(F|C)} = \frac{0,089 \times \mathbb{P}(H)}{0,094 \times \mathbb{P}(F)} = \frac{0,089 \times 1,1}{0,094} = 1,04 > 1.$$

Nous en déduisons que $\mathbb{P}(H|C) > \mathbb{P}(F|C)$. Autrement dit, parmi les chômeurs, les hommes sont plus nombreux que les femmes.

Pour s'entraîner (solutions p. 542)

74.1 Considérons les deux résistances r_1 et r_2 illustrées sur la Figure 74.1, et considérons les deux événements R_i = « la résistance r_i est en panne », avec $i = 1$ ou 2. Nous notons $\mathbb{P}(R_i) = p_i$.

1. Donner la probabilité que le système en parallèle soit en panne, en fonction de p_1 et p_2 . Même question pour le système en série.

2. Calculer la probabilité que la résistance r_1 soit en panne, sachant que le système est en panne.

74.2

Trois symptômes s_1 , s_2 et s_3 correspondent à une maladie donnée. Ces trois symptômes peuvent ou non arriver simultanément. La probabilité d'observer simultanément les trois symptômes est égale à $\frac{1}{8}$. La probabilité d'observer s_1 , s_3 mais pas s_2 est $\frac{1}{4}$; la probabilité d'observer s_2 , s_3 mais pas s_1 est $\frac{1}{8}$; la probabilité d'observer s_1 , s_2 mais pas s_3 est nulle; la probabilité d'observer s_3 mais pas s_1 est $\frac{1}{8}$; la probabilité d'observer s_1 mais pas s_2 est $\frac{3}{8}$; la probabilité d'observer s_2 mais ni s_1 , ni s_3 est $\frac{1}{4}$. Nous notons S_i , pour $i = 1; 2; 3$ l'événement « le symptôme s_i est observé ». Montrer que $\mathbb{P}(S_1 \cap S_2 \cap S_3) = \mathbb{P}(S_1) \times \mathbb{P}(S_2) \times \mathbb{P}(S_3)$ mais que S_1, S_2, S_3 ne sont pas indépendants.

74.3 Dans une université de médecine, le taux de réussite au Numerus Clausus de première année est de 80% pour les redoublants,

et de 30% pour les non redoublants. Une promotion de première année contient 700 non redoublants et 300 redoublants. Quel est le taux de réussite au Numerus Clausus ?

74.4 Dans un campus universitaire se trouvent trois écoles d'ingénieur E_1, E_2, E_3 . Ces écoles ont respectivement 140, 160 et 150 élèves (chaque élève n'est étudiant que d'une seule école). La proportion de filles dans chaque école est respectivement de 10 %, 80 %, 50 %. Quel est le pourcentage de filles dans ce campus ? Léa est étudiante dans ce campus : quelle est la probabilité qu'elle soit élève de l'école E_1 ?

74.5

En France, 90 % des cirrhoses du foie chez l'homme sont dues à la consommation excessive d'alcool. Chez la femme, cette proportion est de 70 %. De plus, environ cinq hommes sur 10 000 et une femme sur 10 000 sont atteints de cirrhose du foie. La proportion de buveurs excessifs est quatre fois plus élevée chez les hommes que chez les femmes. D'après ces chiffres, une femme alcoolique a-t-elle un risque moins élevé de développer une cirrhose du foie qu'un homme alcoolique ?

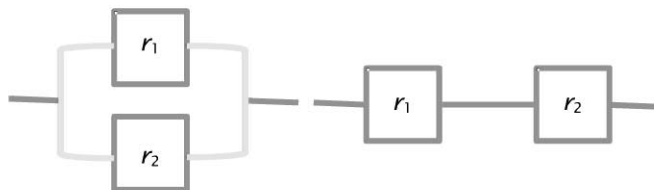


Figure 74.1 – Résistances en parallèle (à gauche), en série (à droite).

Variable aléatoire réelle

Loi d'une variable aléatoire réelle

Le monde de l'aléatoire est défini par des variables, elles-mêmes aléatoires. Nous présentons maintenant la notion de « variable aléatoire ».

1. Variable aléatoire réelle

Définition 75.1

Une **variable aléatoire réelle** X est une application de Ω dans $\mathbb{R}$, $X : \begin{cases} \Omega \longrightarrow \mathbb{R} \\ \omega \longmapsto X(\omega) \end{cases}$.

Pour tout $\omega \in \Omega$, la valeur $X(\omega)$ est une **réalisation** de la variable aléatoire X . L'ensemble $X(\Omega)$ est l'ensemble des **valeurs prises par la variable aléatoire**. Nous utilisons parfois le terme de variable aléatoire à la place de variable aléatoire réelle.

- Si $X(\Omega)$ est un ensemble fini ou dénombrable : la variable aléatoire X est **discrète**.
- Si $X(\Omega)$ est un intervalle ou l'union de plusieurs intervalles de $\mathbb{R}$: la variable aléatoire X est **continue**.

Exemple

Sur l'univers Ω = « ensemble des pièces produites par une machine », définissons les variables aléatoires suivantes :

- pour toute pièce $\omega \in \Omega$, $X(\omega) = 0$ si la pièce est conforme au cahier des charges, $X(\omega) = 1$ sinon. C'est une **variable aléatoire discrète** : $X(\Omega) = \{0; 1\}$.
- Pour toute pièce $\omega \in \Omega$, $Y(\omega)$ est la durée de vie de la pièce ω . C'est une **variable aléatoire continue** : $Y(\Omega) = \mathbb{R}^+$.

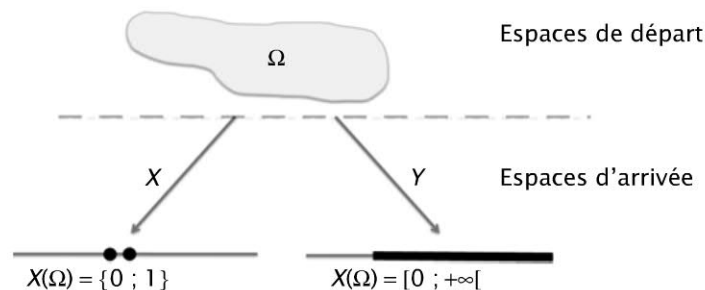


Figure 75.1 – Variables aléatoires discrète et continue.



Point méthode

Avec les variables aléatoires, il n'est plus nécessaire de définir Ω , seul l'ensemble $X(\Omega)$ sera nécessaire.

Définissons maintenant la **loi d'une variable aléatoire réelle**, qu'elle soit discrète ou continue.

2. Loi d'une variable aléatoire réelle discrète

Définition 75.2

Si X est une variable aléatoire **discrète**, la **loi** de X est définie par la famille de réels $(p_X(x))_{x \in X(\Omega)}$ compris entre 0 et 1 vérifiant $\sum_{x \in X(\Omega)} p_X(x) = 1$ et telle que pour tout événement $A \subset X(\Omega)$, la probabilité que X appartienne à A est

$$\mathbb{P}(X \in A) = \sum_{x \in A} p_X(x).$$

Plus précisément, $\forall x \in X(\Omega)$, $p_X(x) = \mathbb{P}(\{\omega \in \Omega; X(\omega) = x\})$, ce que nous noterons plus simplement $\mathbb{P}(X = x)$, c'est-à-dire $p_X(x) = \mathbb{P}(X = x)$.

La loi discrète peut être représentée par un **diagramme en bâtons** (cf. Figure 75.2).

Exemple

Nous lançons un dé non biaisé. Soit X la face qui apparaît lors d'un tirage. X est une **variable aléatoire discrète** à valeurs dans $X(\Omega) = \{1; 2; 3; 4; 5; 6\}$ et puisque le dé est non biaisé, nous avons, pour tout $k \in X(\Omega)$, $\mathbb{P}(X = k) = p_X(k) = \frac{1}{6}$. La loi de X est représentée sur la Figure 75.2. De plus, $\mathbb{P}(X \in \{1; 3; 5\}) = p_X(1) + p_X(3) + p_X(5) = \frac{3}{6} = \frac{1}{2}$. Nous verrons à la Fiche 81 que c'est une loi discrète uniforme.

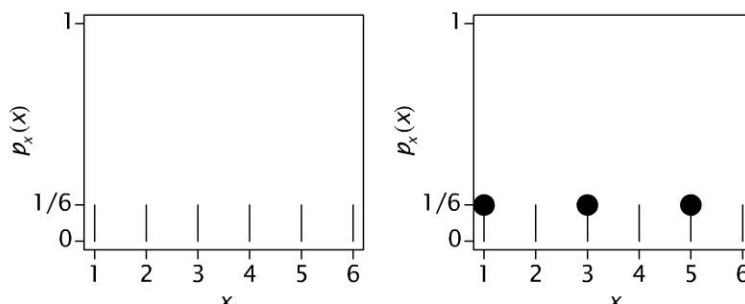


Figure 75.2 – Variable aléatoire discrète : diagramme en bâtons à gauche, et

$$\mathbb{P}(X \in \{1; 3; 5\}) = p_X(1) + p_X(3) + p_X(5) = \frac{3}{6} = \frac{1}{2} \text{ à droite.}$$

3. Loi d'une variable aléatoire réelle continue

Définition 75.3

Soit X une variable aléatoire **continue**. Nous disons que la **loi de X admet une densité de probabilité sur $\mathbb{R}$** s'il existe une densité de probabilité f_X définie sur $\mathbb{R}$, telle que pour tout événement $A \subset \mathbb{R}$, la probabilité que X appartienne à A est

$$\mathbb{P}(X \in A) = \int_A f_X(x) dx.$$

La loi de X est appelée **loi de probabilité de densité f_X** .

Exemple

Soit X une variable aléatoire continue, à valeurs dans $\mathbb{R}$, et de densité $f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$, représentée sur la Figure 75.3. Nous verrons à la Fiche 83 que c'est une loi gaussienne centrée et réduite. La probabilité que X soit négative est égale à $\mathbb{P}(X \leq 0) = \int_{-\infty}^0 f_X(x) dx = 0,5$. Cette probabilité est l'aire grisée du graphique de droite de la Figure 75.3 : c'est l'aire sous la densité, entre les abscisses $-\infty$ et 0.

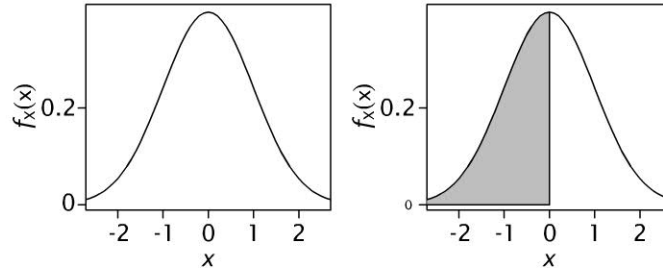


Figure 75.3 – Variable aléatoire continue : densité à gauche, calcul de $\mathbb{P}(X \leq 0)$ à droite.



- Soit X une variable aléatoire continue de densité f_X . Les fortes valeurs de densité ont une grande probabilité d'apparaître (par exemple, dans le cas de la Figure 75.3, les valeurs comprises entre -1 et $+1$), tandis que les faibles valeurs de densité ont une faible probabilité d'apparaître (par exemple, dans le cas de la Figure 75.3, les valeurs supérieures à $+2,5$ ou inférieures à $-2,5$).
- Soit X une variable aléatoire continue de densité f_X . Pour tout $a \in \mathbb{R}$, la probabilité $\mathbb{P}(X = a)$ est nulle, car $\mathbb{P}(X = a) = \int_{\{a\}} f_X(x) dx = 0$.
- Une variable aléatoire continue peut ne pas avoir de densité, c'est le cas appelé « mixte », un mélange entre loi discrète et loi à densité. Nous verrons un exemple dans ce livre, lors de l'Application de la Fiche 80. Dans ce cas, la probabilité $\mathbb{P}(X = a)$ n'est plus nécessairement nulle.

Ne pas confondre loi et variable aléatoire : si nous lançons deux dés équilibrés, soit X le résultat du premier dé et Y celui du deuxième, alors X et Y ont clairement la même loi, mais ne sont pas égaux.

Application Un exemple de loi

Nous tirons un nombre réel au hasard entre 0 et 2, et nous notons X la variable aléatoire associée à cette expérience aléatoire. Quelle est la loi de X ?

X est à valeurs dans l'intervalle $[0; 2]$, c'est-à-dire $X(\Omega) = [0; 2]$. X est une **variable aléatoire continue**, de densité f_X à déterminer. Faisons les remarques suivantes :

1. Nous sommes dans le cadre de la Définition 73.5 de la Fiche 73 :

pour tout $A \subset [0; 2]$, $\mathbb{P}(X \in A) = \frac{\ell(A)}{\ell([0; 2])} = \frac{\ell(A)}{2}$ où $\ell(A)$ désigne la longueur de l'événement A .

2. De plus, d'après la Définition 75.3, pour tout ensemble A de $\mathbb{R}$, nous avons $\mathbb{P}(X \in A) = \int_A f_X(x) dx$.

Nous remarquons que si A est inclus dans $\mathbb{R}_-^* \cup]2; +\infty[$, nous avons $\mathbb{P}(X \in A) = 0$. Ceci nous donne une indication sur la densité f_X : f_X est nulle sur $\mathbb{R}_-^* \cup]2; +\infty[$.

Combinant les deux remarques précédentes, f_X est de la forme $f_X(x) = C \mathbb{1}_{x \in [0;2]}$, où C est une constante réelle à déterminer et $\mathbb{1}_{x \in [0;2]}$ vaut 1 si $x \in [0;2]$ et 0 si $x \notin [0;2]$.

Maintenant, f_X devant être une densité, nous avons $\int_{\mathbb{R}} f_X(x) dx = 1$, ce qui détermine C de manière unique. Nous trouvons ainsi $C = \frac{1}{2}$.

Nous concluons que la densité de la variable X est $f_X(x) = \frac{1}{2} \mathbb{1}_{x \in [0;2]}$. Nous verrons à la Fiche 84 que X suit la loi uniforme sur $[0;2]$.

Pour s'entraîner (solutions p. 543)

75.1 Au casino, la roulette comprend 37 numéros allant de 0 à 36, et n'est pas truquée : chaque numéro a la même probabilité de sortir. M. Untel mise 1 euro sur les nombres impairs : si un numéro impair sort, il rempoche son euro et gagne un euro supplémentaire (le 0 n'est ni pair ni impair), sinon, il perd son euro. Notons G le gain de M. Untel. Quelle est la loi de G ?

75.2 Trois dés sont jetés, et la somme des résultats apparus sur les faces est notée S . Que vaut $S(\Omega)$? Lequel de ces événements $\{S = 7\}$ et $\{S = 8\}$ est le plus probable ?

75.3  [www](#)

Soit X le salaire mensuel net à la sortie d'une formation, en 2009, et $f_X(x) = A e^{-[x-1500]} \mathbb{1}_{x \geq 0}$ sa densité de probabilité.

1. Que vaut la constante A ?
2. Tracer l'allure de courbe de la densité. Quelle est l'aire sous cette courbe ?
3. Géométriquement, comment calculer la probabilité d'avoir un salaire inférieur à 1500 euros ? Donner la valeur de cette probabilité.
4. De même, comment calculer la probabilité d'avoir un salaire supérieur à 10 000 euros par mois ?
5. À la sortie de cette formation, quel est le salaire perçu par plus de 95% des sala-

riés ? Représenter cette valeur sur le schéma. Quelle est l'aire sous la courbe entre 0 et cette valeur ?

75.4 Soit une source isotrope émettant des particules dans le plan, et positionnée aux coordonnées $(0;1)$ du plan. Notons X la position sur l'axe des abscisses de l'impact d'une particule envoyée par la source. X est une variable aléatoire de densité définie pour tout $x \in \mathbb{R}$ par $f_X(x) = \frac{A}{1+x^2}$.

1. Que vaut la constante A ?
2. Quelle est la probabilité que X soit compris entre -1 et $+1$?

75.5 Pour un alunissage de fusée, considérons l'écart X (en mètres) entre la position d'alunissage de la fusée et la position prévue par les astronautes. Supposons que X est une variable de densité $f_X(x) = A x e^{-h^2 x^2}$ sur $\mathbb{R}^+$ et nulle sur $\mathbb{R}^-$, où $h > 0$ est un paramètre réel fixé.

1. Exprimer la constante A en fonction de h .
2. Quelle est la probabilité que la fusée se pose à moins de 10 m de la cible ? Dans 90 % des cas, la fusée se pose à moins de 10 m de la position prévue par les astronautes. En déduire la valeur du paramètre h .

Fonction de répartition

Rappelons que la loi d'une variable aléatoire réelle X est donnée par la famille de réels $(p_X(x))_{x \in X(\Omega)}$ si X est discrète, ou par la densité $x \mapsto f_X(x)$ définie sur $\mathbb{R}$ si X est à densité.

Une autre manière de caractériser la loi d'une variable aléatoire réelle X est de donner sa **fonction de répartition**, notée F_X .

1. Fonction de répartition d'une variable aléatoire réelle

Définition 76.1

Soit X une variable aléatoire réelle. Sa **fonction de répartition** est la fonction F_X telle que :

$$F_X : \begin{cases} \mathbb{R} \longrightarrow [0; 1] \\ x \longmapsto F_X(x) = \mathbb{P}(X \leq x) \end{cases}$$

Voici ses propriétés générales :

Théorème 76.1

Soit F_X une fonction de répartition, alors :

- F_X est une fonction croissante sur $\mathbb{R}$, et

$$\lim_{x \rightarrow -\infty} F_X(x) = 0, \quad \lim_{x \rightarrow +\infty} F_X(x) = 1.$$

- Pour tous réels $a < b$,

$$\mathbb{P}(a < X \leq b) = F_X(b) - F_X(a),$$

et

$$\mathbb{P}(a \leq X \leq b) = F_X(b) - F_X(a) + \mathbb{P}(X = a).$$

Rappelons que si la variable aléatoire X est à densité, alors $\mathbb{P}(X = a) = 0$.

2. Cas particulier des variables aléatoires discrètes

Supposons que la variable aléatoire X est discrète à valeurs dans $X(\Omega) = \{x_0 < x_1 < \dots < x_n < \dots\}$. Notons $(p_X(x_i))_{x_i \in X(\Omega)}$ sa loi. Nous avons les propriétés suivantes :

- Les valeurs de F_X sont les sommes cumulées des valeurs de $(p_X(x_i))_{x_i \in X(\Omega)}$:

$$\forall x \in \mathbb{R}, F_X(x) = \sum_{x_i \leq x} p_X(x_i).$$

- Inversement, nous retrouvons la loi $(p_X(x_i))_{x_i \in X(\Omega)}$ à partir de F_X :

- $p_X(x_0) = F_X(x_0)$
- $\forall x_i \neq x_0, p_X(x_i) = F_X(x_i) - F_X(x_{i-1})$

- F_X est une fonction en « escaliers » : elle est discontinue aux points $x_i \in X(\Omega)$. Elle est continue à droite et admet une limite à gauche en tout point de $\mathbb{R}$ (nous disons parfois càdlàg).



Point méthode

Pour donner la loi d'une variable aléatoire discrète, il est équivalent de spécifier $(p_X(x_i))_{x_i \in X(\Omega)}$ ou la fonction de répartition F_X , puisque l'une se déduit de l'autre.

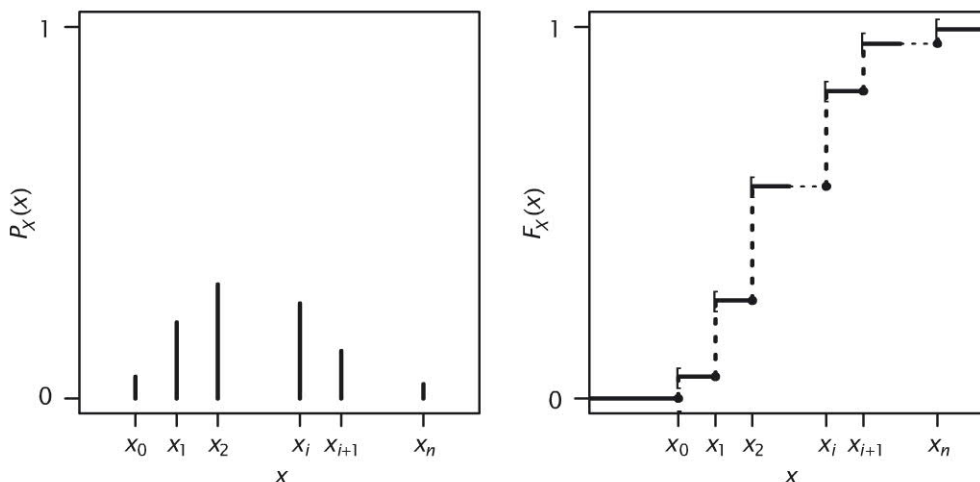


Figure 76.1 – Graphes de la loi $(p_X(x_i))_{x_i \in X(\Omega)}$ (à gauche) et de la fonction de répartition F_X (à droite).

3. Cas particulier des variables aléatoires à densité

Supposons que la variable aléatoire X admet une densité f_X . Nous avons les deux propriétés suivantes :

- La fonction de répartition s'écrit :

$$\forall x \in \mathbb{R}, \quad F_X(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^x f_X(t) dt.$$

- Pour tous réels $a < b$,

$$\mathbb{P}(X \in [a; b]) = \mathbb{P}(X \in]a; b]) = \mathbb{P}(X \in [a; b[) = \mathbb{P}(X \in]a; b[) = F_X(b) - F_X(a) = \int_a^b f_X(t) dt.$$

La Figure 76.2 illustre cette propriété.



Point méthode

Pour donner la loi d'une variable aléatoire admettant une densité, il est équivalent de spécifier la densité f_X ou la fonction de répartition F_X , puisque l'une se déduit de l'autre : F_X est la primitive de f_X , et f_X est la dérivée de F_X en tout point où F_X est dérivable.

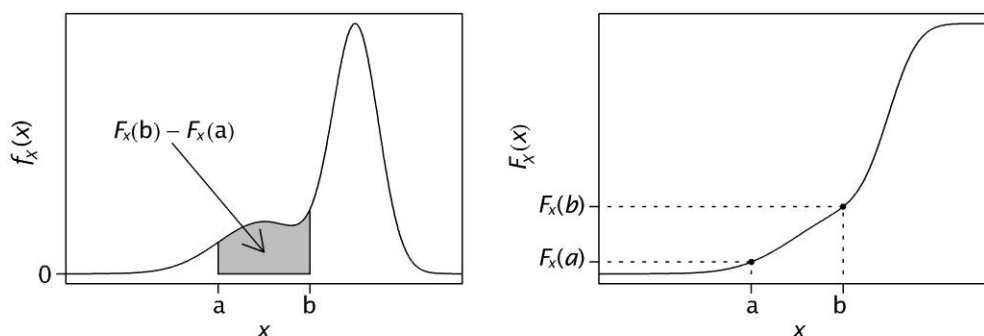


Figure 76.2 – Graphes de la densité f_X (à gauche) et de la fonction de répartition F_X (à droite).

Considérons la nappe phréatique d'Alsace, autour de Colmar. Cet endroit est constitué de 4 types de géologie. Pour chaque type $i = 1; 2; 3$ et 4, la nappe phréatique a une profondeur X_i de densité f_{X_i} . La géologie de type 4 couvre 85 % de la nappe considérée, et chacun des autres types couvrent 5 % de la nappe. Un forage de 12 mètres est réalisé dans cette zone. Soit X la profondeur de l'eau dans le forage. Quelle est la probabilité de trouver de l'eau dans le forage ? Quelle est la loi de X ?

Notons A_i pour $i = 1; 2; 3$ et 4, l'événement $A_i = \ll$ la nappe est de type $i \gg$. L'énoncé nous donne $\mathbb{P}(A_i) = 0,05$ pour $i = 1; 2; 3$ et $\mathbb{P}(A_4) = 0,85$.



Nous ne savons pas dans quel type de géologie le forage est réalisé. Nous allons donc prendre en compte les 4 types de géologie, grâce à la formule des probabilités totales.

- La probabilité de trouver de l'eau dans le forage est $\mathbb{P}(X \leq 12)$.
- Si le forage est réalisé dans la géologie de type 1, nous avons $\mathbb{P}(X_1 \leq 12)$. Ce qui s'écrit aussi $\mathbb{P}(X \leq 12 | A_1) = \mathbb{P}(X_1 \leq 12) = \int_{-\infty}^{12} f_{X_1}(t) dt$.
- A_1, A_2, A_3, A_4 forment une partition de Ω , où Ω est la nappe phréatique d'Alsace, autour de Colmar.
- La formule des probabilités totales donne $\mathbb{P}(X \leq 12) = \sum_{i=1}^4 \mathbb{P}(X \leq 12 | A_i) \times \mathbb{P}(A_i)$. Ce qui s'écrit aussi

$$\mathbb{P}(X \leq 12) = \sum_{i=1}^4 \mathbb{P}(X_i \leq 12) \times \mathbb{P}(A_i) = 0,05 \times \left(\int_{-\infty}^{12} f_{X_1}(t) + f_{X_2}(t) + f_{X_3}(t) dt \right) + 0,85 \times \int_{-\infty}^{12} f_{X_4}(t) dt.$$

Les mêmes calculs permettent d'obtenir $F_X(x) = \mathbb{P}(X \leq x)$:

$$F_X(x) = \mathbb{P}(X \leq x) = 0,05 \times \left(\int_{-\infty}^x f_{X_1}(t) + f_{X_2}(t) + f_{X_3}(t) dt \right) + 0,85 \times \int_{-\infty}^x f_{X_4}(t) dt,$$

ou encore

$$F_X(x) = 0,05 \times \sum_{i=1}^3 F_{X_i}(x) + 0,85 \times F_{X_4}(x),$$

ce qui nous définit la loi de X .



Enfin, à titre d'information, nous donnons la densité de X . Il suffit de dériver la fonction de répartition F_X :

$$f_X(x) = F'_X(x) = 0,05 \times \sum_{i=1}^3 f_{X_i}(x) + 0,85 \times f_{X_4}(x).$$

Pour s'entraîner (solutions p. 543)

76.1 Soit X le résultat du lancer d'un dé équilibré. Calculer et tracer la fonction de répartition de X .

76.2 Soit la variable aléatoire X de fonction de répartition définie par $F_X(t) = 0$ si $t < 3$; $F_X(t) = 0,3$ si $t \in [3; 5[$; $F_X(t) = 0,6$ si $t \in [5; 8[$ et $F_X(t) = 1$ si $t \geq 8$.

1. Tracer F_X . Soit $X(\Omega)$ l'ensemble des valeurs prises par X . Préciser $X(\Omega)$ et pour tout $x \in X(\Omega)$, donner $\mathbb{P}(X = x)$.

Calculer :

2. $\mathbb{P}(X < 2)$;

3. $\mathbb{P}(X > \frac{3}{2})$;

4. $\mathbb{P}(3 < X < 5)$;

5. $\mathbb{P}(3 < X \leq 5)$;

6. $\mathbb{P}(3 \leq X < 5)$;

7. $\mathbb{P}(3 \leq X \leq 5)$.

76.3 Soit X le nombre de forages aboutissant à du pétrole parmi 12 forages effectués. Supposons que la loi de X est donnée par $\mathbb{P}(X = k) = p_k$ pour tout $k \in \{0; \dots; 12\}$.

1. Exprimer la fonction de répartition de X en fonction des p_k , pour $k \in \{0; \dots; 12\}$.

2. Pour éviter la banqueroute, il faut que trois forages au moins donnent du pétrole. Exprimer la probabilité de cet événement en fonction de la fonction de répartition de X .

76.4 Soit X une variable aléatoire représentant la durée de vie d'une batterie, et $F_X(t) = (1 - e^{-\lambda t}) \mathbb{1}_{t \geq 0}$ sa fonction de répartition, où $\lambda > 0$ est un paramètre fixé.

1. Tracer F_X .

2. Le constructeur indique que dans 90 % des cas, sa batterie dure au moins 10 unités de temps. Que vaut λ ?

3. Quelle est la probabilité que la batterie dure plus de 5 unités de temps ?

4. Sachant que la batterie a déjà duré plus de 10 unités de temps, quelle est la probabilité qu'elle dure encore 5 unités de temps supplémentaires ? Commenter ce résultat.

5. Montrer que X admet une densité et la déterminer.

76.5 Nous reprenons le dernier énoncé de la Fiche 75, où X désigne l'écart (en mètres) entre la position d'alunissage de la fusée et la position prévue par les astronautes. Calculer la fonction de répartition de X .

Par définition, une variable aléatoire « varie ». Nous pouvons définir la valeur moyenne d'une variable aléatoire : cette valeur est son espérance.

1. Espérance d'une variable aléatoire réelle

L'**espérance** d'une variable aléatoire réelle X est la valeur attendue en moyenne, pour X . C'est donc **la moyenne pondérée** des valeurs que X peut prendre, les poids étant les probabilités que ces valeurs soient prises.

Définition 77.1

Soit X une variable aléatoire réelle. L'**espérance** de X , notée $\mathbb{E}(X)$, est définie par :

$$\mathbb{E}(X) = \sum_{x \in X(\Omega)} xp_X(x) \quad \text{si } X \text{ est discrète,} \quad (77.1)$$

$$\mathbb{E}(X) = \int_{-\infty}^{+\infty} xf_X(x)dx \quad \text{si } X \text{ est à densité,} \quad (77.2)$$

où, si X est discrète, $(p_X(x))_{x \in X(\Omega)}$ désigne sa loi, et si X est à densité, f_X désigne sa densité.



- Si $\mathbb{E}(X) = 0$, nous disons que la variable aléatoire X est **centrée**.
- La définition de l'espérance suppose que les quantités (77.1) et (77.2) existent. Pour cela, nous aurons besoin des notions vues aux Fiches 46 et 53.

Par exemple, si X a pour densité définie pour tout $x \in \mathbb{R}$ par $f_X(x) = \frac{1}{\pi} \times \frac{1}{1+x^2}$ (cette densité est appelée densité de Cauchy), la quantité (77.2) n'est pas définie (voir la Fiche 53). Nous disons que la variable X n'a pas d'espérance.

- Dans le cas d'une variable aléatoire X discrète, l'espérance (si elle existe) est un réel qui n'est pas nécessairement inclus dans $X(\Omega)$.
- Par contre, dans le cas d'une variable aléatoire à densité, $X(\Omega)$ est un intervalle de $\mathbb{R}$, et l'espérance (si elle existe) est incluse dans $X(\Omega)$.

Exemple

Dans le cas discret, imaginons des poids $p(x_i)$ répartis sur une barre (de poids nul), aux abscisses x_i , et tels que $\sum_i p(x_i) = 1$. Le centre de gravité de ce montage est défini par le point μ où la barre peut être posée en gardant l'équilibre. Ce point de gravité n'est autre que $\sum_i x_i p(x_i)$, qui vaut 0,61 dans l'exemple de la Figure 77.1. Le centre de gravité μ est l'analogue de l'espérance de la variable aléatoire X prenant les valeurs x_i avec probabilités $p(x_i)$.



Figure 77.1 – Avec les poids $p(0) = 0,51$, $p(1) = 0,38$, $p(2) = 0,10$, $p(3) = 0,01$, placés aux abscisses 0, 1, 2 et 3, le centre de gravité est à l'abscisse 0,61.

Dans le cas continu, imaginons que la barre a une densité de masse, valant $f(x)$ à chaque abscisse x et telle que $\int_{\text{barre}} f(x)dx = 1$. Alors le centre de masse de la barre n'est autre que $\int_{\text{barre}} xf(x)dx$. Le centre de masse de la barre est l'analogue de l'espérance de la variable aléatoire X ayant pour densité f .

Théorème 77.1 Propriétés de l'espérance

- L'espérance d'une constante est la constante : soit a un réel, alors

$$\mathbb{E}(a) = a.$$

- L'espérance est linéaire : soient a, b deux réels, et X, Y deux variables aléatoires réelles, alors

$$\mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y).$$

- Si X et Y sont deux variables aléatoires réelles de même loi, elles ont la même espérance.
- Si X et Y sont deux variables aléatoires réelles telles que $X \leq Y$ alors $\mathbb{E}(X) \leq \mathbb{E}(Y)$.

2. Espérance d'une fonction d'une variable aléatoire réelle

Reprenons l'exemple de la Figure 77.1 avec la variable aléatoire discrète X prenant les valeurs 0, 1, 2 et 3 avec probabilités $p(0)$, $p(1)$, $p(2)$ et $p(3)$. Soit g une fonction numérique définie sur $\mathbb{R}$. Alors $g(X)$ est une nouvelle variable aléatoire, prenant les valeurs $g(0)$, $g(1)$, $g(2)$ et $g(3)$ avec probabilités $p(0)$, $p(1)$, $p(2)$ et $p(3)$. Son espérance est donc

$$\mathbb{E}(g(X)) = \sum_{i=0}^3 g(i)p(i).$$

Donnons maintenant la définition exacte :

Définition 77.2

Soit X une variable aléatoire réelle, et g une fonction numérique définie sur $\mathbb{R}$. L'**espérance** de $g(X)$, notée $\mathbb{E}(g(X))$ est définie par :

$$\begin{aligned} \mathbb{E}(g(X)) &= \sum_{x \in X(\Omega)} g(x)p_X(x) && \text{si } X \text{ est discrète,} \\ \mathbb{E}(g(X)) &= \int_{-\infty}^{+\infty} g(x)f_X(x)dx && \text{si } X \text{ est à densité} \end{aligned}$$

où, si X est discrète, $(p_X(x))_{x \in X(\Omega)}$ désigne sa loi, et si X est à densité, f_X désigne sa densité.



- La fonction g ne change pas les « poids » $p_X(x)$ (pour le cas discret), et f_X (pour le cas densité). Elle n'intervient que sur les abscisses x .
- Il faut noter que $\mathbb{E}(g(X))$ n'est en général pas égal à $g(\mathbb{E}(X))$. Par exemple, *sauf si X est constante*, nous avons $\mathbb{E}(X^2) \neq (\mathbb{E}(X))^2$.

Application **Le boulanger**

Un boulanger a remarqué que le nombre de pains X qui lui sont demandés chaque jour suit une loi de densité f_X , nulle sur $\mathbb{R}^-$. La vente d'un pain rapporte un bénéfice net de β euros, chaque pain invendu en fin de journée est un déficit de δ euros. Le boulanger cherche à savoir combien il doit préparer de pains pour qu'en moyenne, son gain journalier soit maximal. Répondre à ce boulanger.



Remarquons que y est un entier non aléatoire (c'est le nombre de pains que le boulanger a choisi de faire), et G est aléatoire (car il dépend du nombre de demandes, ce nombre étant aléatoire). Remarquons également que $\mathbb{E}(G)$ est le gain journalier moyen du boulanger. L'exercice nous demande de chercher la valeur de y qui maximise $\mathbb{E}(G)$.

Exprimons G le gain journalier du boulanger en fonction de X, β, δ et y .



Nous distinguons deux cas. Soit le boulanger vend tous les pains qu'il a fabriqués, ce cas comprend aussi la situation où le boulanger ne satisfait pas toutes les commandes : c'est le cas $X > y$; soit il lui reste des invendus en fin de journée : c'est le cas $X \leq y$. Dans le premier cas, le boulanger a un bénéfice net de βy . Dans le second cas, il a un bénéfice de βX et une perte de $(y - X)\delta$.

Nous pouvons donc écrire $G = \beta y \mathbb{1}_{X > y} + (\beta X - (y - X)\delta) \mathbb{1}_{X \leq y} = \beta y \mathbb{1}_{X > y} + ((\beta + \delta)X - y\delta) \mathbb{1}_{X \leq y}$, où nous rappelons que $\mathbb{1}_{X \in A}$ est égal à 1 si $X \in A$ et 0 si $X \notin A$.

Calculons maintenant le gain moyen journalier :

$$\begin{aligned} \mathbb{E}(G) &= \mathbb{E}[\beta y \mathbb{1}_{X > y} + ((\beta + \delta)X - y\delta) \mathbb{1}_{X \leq y}] \\ &= \beta y \times \mathbb{E}(\mathbb{1}_{X > y}) + (\beta + \delta) \times \mathbb{E}(X \mathbb{1}_{X \leq y}) - \delta y \times \mathbb{E}(\mathbb{1}_{X \leq y}), \end{aligned}$$

par la propriété de linéarité de l'espérance.

Calculons chacun de ces trois termes.

1. Remarquons que le premier terme $\mathbb{1}_{X > y}$ vaut 1 si $X > y$, ce qui arrive avec probabilité $\mathbb{P}(X > y)$ et 0 sinon, ce qui arrive avec probabilité $\mathbb{P}(X \leq y)$. Donc

$$\mathbb{E}(\mathbb{1}_{X > y}) = 1 \times \mathbb{P}(X > y) + 0 \times \mathbb{P}(X \leq y) = \mathbb{P}(X > y) = \int_y^{+\infty} f_X(x) dx = 1 - \int_0^y f_X(x) dx.$$

2. Pour le second terme, nous avons $\mathbb{E}(X \mathbb{1}_{X \leq y}) = \int_{\mathbb{R}} x \mathbb{1}_{x \leq y} f_X(x) dx = \int_0^y x f_X(x) dx$.

3. Enfin, le troisième terme est $\mathbb{E}(\mathbb{1}_{X \leq y}) = \mathbb{P}(X \leq y) = \int_0^y f_X(x) dx$ par le même raisonnement que le premier terme.

Finalement, $\mathbb{E}(G) = \beta y \times \left(1 - \int_0^y f_X(x) dx\right) + (\beta + \delta) \times \int_0^y x f_X(x) dx - \delta y \times \int_0^y f_X(x) dx$.

Nous obtenons donc $\mathbb{E}(G)$ comme une fonction de y , écrivons donc $\mathbb{E}(G) = g(y)$. La recherche du maximum de $\mathbb{E}(G) = g(y)$ est alors un problème classique d'optimisation. Résolvons $g'(y) = 0$. Nous avons

$$\begin{aligned} \frac{d}{dy} \mathbb{E}(G) = g'(y) &= \beta \times \left(1 - \int_0^y f_X(x) dx\right) - \beta y \times f_X(y) + (\beta + \delta) \times y f_X(y) - \delta \\ &\quad \times \int_0^y f_X(x) dx - \delta y \times f_X(y) \\ &= \beta \times \left(1 - \int_0^y f_X(x) dx\right) - \delta \times \int_0^y f_X(x) dx. \end{aligned}$$

La résolution de $g'(y) = 0$ entraîne $\int_0^y f_X(x) dx = \frac{\beta}{\beta + \delta}$, c'est-à-dire $F_X(y) = \frac{\beta}{\beta + \delta}$, où F_X est la fonction de répartition de X . La quantité optimale de pains à préparer par le boulanger est donc la solution y_0 de

$$F_X(y_0) = \frac{\beta}{\beta + \delta}.$$



- y_0 est bien un maximum local de la fonction g car $g''(y_0) = -(\beta + \delta) f_X(y_0) \leq 0$.
- Nous verrons à la Fiche 79 que y_0 s'appelle **quantile d'ordre $\frac{\beta}{\beta + \delta}$ de la loi de X** . Suivant la loi de X , remarquons que ce quantile ne sera peut-être pas un nombre entier, dans ce cas, il suffira de prendre la partie entière.

Fiche 38

Fiche 79

Pour s'entraîner (solutions p. 544)

77.1 Soient $a \in \mathbb{R}$ et X une variable aléatoire réelle constante égale à a . Montrer que $\mathbb{E}(X) = a$.

77.2 Un signal X émet des fréquences f_1 , f_2 ou f_3 avec probabilités $\frac{1}{2}$, $\frac{1}{8}$ et $\frac{3}{8}$. Quelle est la fréquence moyenne émise ?

77.3 Soit X une variable aléatoire et A un ensemble de $\mathbb{R}$. Montrer que $\mathbb{E}(1_{X \in A}) = \mathbb{P}(X \in A)$, où $1_{X \in A} = 1$ si $X \in A$, et $1_{X \in A} = 0$ si $X \notin A$.

77.4 M. Untel achète une grille d'un jeu à gratter. Le prix de la grille est 2 euros.

La probabilité qu'il gagne est $p \in]0; 1[$. S'il gagne, il obtient une cagnotte de c euros. Que doivent vérifier p et c pour que l'organisme du jeu soit assuré d'avoir un gain moyen strictement positif ? Que doit être la cagnotte maximale autorisée par l'organisme du jeu si $p = \frac{1}{2}$? Si $p = \frac{1}{10}$? Si $p = \frac{1}{10^6}$?

77.5 Considérons l'exercice d'application de la Fiche 76. Calculer la hauteur d'eau moyenne de la nappe phréatique dans la région étudiée, en fonction des hauteurs d'eau moyennes des quatre types de géologie présentes.

Moments, Variance et Écart-type

Nous avons vu que l'espérance d'une variable aléatoire réelle X décrit la valeur moyenne de X . Toutefois, l'espérance ne donne aucune indication quant au caractère dispersé ou non de la variable aléatoire X . Cette indication sera donnée par la variance, ou encore par l'écart-type de la variable aléatoire X .



1. Moments d'une variable aléatoire réelle

Définition 78.1

Soit $r \in \mathbb{N}^*$. Le **moment d'ordre r** de la variable aléatoire réelle X est le réel défini par $\mu_r = \mathbb{E}(X^r)$. Le **moment centré d'ordre r** est le réel défini par $\tilde{\mu}_r = \mathbb{E}((X - \mathbb{E}(X))^r)$.



- Le moment d'ordre 1 est l'espérance $\mathbb{E}(X)$, définie à la Fiche 77.
- Le moment centré d'ordre 1 est nul : $\tilde{\mu}_1 = \mathbb{E}(X - \mathbb{E}(X)) = \mathbb{E}(X) - \mathbb{E}(X) = 0$.
- Si deux variables aléatoires X et Y ont la même loi, alors elles ont les mêmes moments.

2. Variance et Écart-type d'une variable aléatoire réelle

► Approche intuitive

Soit une variable aléatoire réelle X d'espérance $\mathbb{E}(X)$. La **variance** de X est une mesure de la dispersion des valeurs prises par X , autour de son espérance. C'est donc une information de plus.

Définition 78.2

Soit X une variable aléatoire réelle. La **variance** de X , notée $\text{Var}(X)$ est le réel positif défini par le moment centré d'ordre 2 :

$$\text{Var}(X) = \mathbb{E}((X - \mathbb{E}(X))^2),$$

c'est-à-dire :

$$\text{Var}(X) = \sum_{x \in X(\Omega)} (x - \mathbb{E}(X))^2 p_X(x) \quad \text{si } X \text{ est discrète,} \quad (78.1)$$

$$\text{Var}(X) = \int_{-\infty}^{+\infty} (x - \mathbb{E}(X))^2 f_X(x) dx \quad \text{si } X \text{ est à densité,} \quad (78.2)$$

où, si X est discrète, $(p_X(x))_{x \in X(\Omega)}$ désigne sa loi, et si X est à densité, f_X désigne sa densité.

L'**écart-type** de X est la valeur réelle positive $\sqrt{\text{Var}(X)}$.



- Si $\text{Var}(X) = 1$, nous disons que la variable aléatoire X est **réduite**.
- Les définitions de la variance et de l'écart-type supposent que les quantités (78.1) et (78.2) existent. Pour cela, nous aurons besoin des notions vues aux Fiches 46 et 53.
- Notons que l'**écart-type** a la même unité physique que X . L'unité de la **variance** est le carré de l'unité de X .

► Analogie avec la notion de moment d'inertie

Reprenons l'exemple de la barre sans poids de la Figure 77.1, sur laquelle nous avons posé des poids aux abscisses x_i . Le moment d'inertie de cet ensemble de poids autour de son centre de gravité μ est l'analogue de la variance de la variable aléatoire X prenant les valeurs x_i avec probabilités $p(x_i)$, c'est-à-dire $\mathbb{V}\text{ar}(X) = \sum_i (x_i - \mu)^2 p(x_i)$.

Dans le cas continu, imaginons que la barre a une densité de masse, valant $f(x)$ à chaque abscisse x . Alors le moment d'inertie de la barre autour de son centre de masse μ est l'analogue de la variance de la variable aléatoire X ayant pour densité f , c'est-à-dire $\mathbb{V}\text{ar}(X) = \int_{\text{barre}} (x - \mu)^2 f(x) dx$.



La variance caractérise la dispersion de la variable aléatoire. La Figure 78.1 présente des variables aléatoires, de même moyenne, mais de variances différentes : plus la variance σ^2 est grande, plus la dispersion autour de la moyenne est importante.

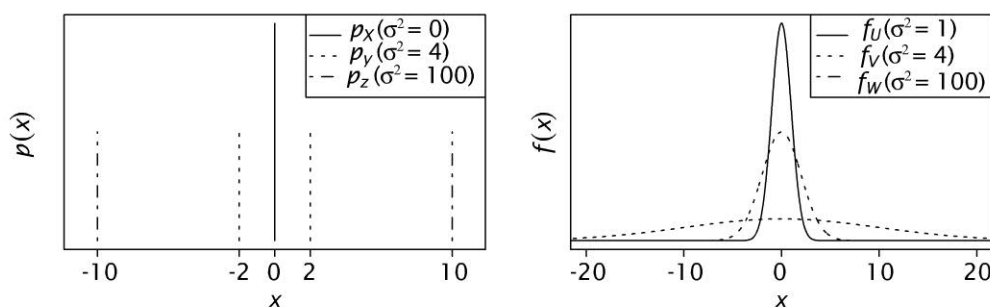


Figure 78.1 – Loïs de variables aléatoires centrées, de diverses variances (cas discret à gauche, continu à droite).

Théorème 78.1 Propriétés de la variance

Soit X une variable aléatoire réelle.

- Par définition, la variance est positive ou nulle : $\mathbb{V}\text{ar}(X) \geq 0$.
- Nous avons $\mathbb{V}\text{ar}(X) = \mathbb{E}((X - \mathbb{E}(X))^2) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$. Ceci s'appelle aussi *Théorème de Huygens*.
- La variance d'une constante est nulle : soit a un réel, alors $\mathbb{V}\text{ar}(a) = 0$.
Réciproquement, le cas $\mathbb{V}\text{ar}(X) = 0$ signifie que la variable aléatoire X n'a aucune dispersion autour de sa moyenne. X est donc une constante.
- Soient a et b deux réels, alors

$$\mathbb{V}\text{ar}(aX + b) = a^2 \mathbb{V}\text{ar}(X).$$

- Si $\mathbb{E}(X) = \mu$ et $\mathbb{V}\text{ar}(X) = \sigma^2$, alors la variable aléatoire $(X - \mu)/\sigma$ est **centrée et réduite** : $\mathbb{E}\left(\frac{X - \mu}{\sigma}\right) = 0$ et $\mathbb{V}\text{ar}\left(\frac{X - \mu}{\sigma}\right) = 1$.
- Si X et Y sont deux variables aléatoires réelles de même loi, alors elles ont la même variance, et donc le même écart-type.

3. Variance d'une fonction d'une variable aléatoire réelle

Reprenons l'exemple de la Figure 77.1 avec la variable aléatoire discrète X prenant les valeurs 0, 1, 2 et 3 avec probabilités $p(0)$, $p(1)$, $p(2)$ et $p(3)$. Soit g une fonction numérique définie sur $\mathbb{R}$. Alors $g(X)$ est une nouvelle variable aléatoire, prenant les valeurs

$g(0), g(1), g(2)$ et $g(3)$ avec probabilités $p(0), p(1), p(2)$ et $p(3)$. Sa variance est donc

$$\mathbb{V}\text{ar}(g(X)) = \sum_{i=0}^3 (g(i) - \mathbb{E}(g(X)))^2 p(i) \quad \text{où} \quad \mathbb{E}(g(X)) = \sum_{i=0}^3 g(i)p(i).$$

Donnons maintenant la définition exacte :

Définition 78.3

Soit X une variable aléatoire réelle, et g une fonction numérique définie sur $\mathbb{R}$. La **variance de $g(X)$** est définie par :

$$\begin{aligned} \mathbb{V}\text{ar}(g(X)) &= \sum_{x \in X(\Omega)} [g(x) - \mathbb{E}(g(X))]^2 p_X(x) \quad \text{si } X \text{ est discrète,} \\ \mathbb{V}\text{ar}(g(X)) &= \int_{-\infty}^{+\infty} [g(x) - \mathbb{E}(g(X))]^2 f_X(x) dx \quad \text{si } X \text{ est à densité,} \end{aligned}$$

où, si X est discrète, $(p_X(x))_{x \in X(\Omega)}$ désigne sa loi, et si X est à densité, f_X désigne sa densité. $\mathbb{E}(g(X))$ désigne l'espérance de $g(X)$.

Application Un alunissage de fusée

Pour un alunissage de fusée, considérons la distance R (en mètres) entre la position d'alunissage de la fusée et la position prévue par les astronautes. Supposons que R soit une variable aléatoire de densité $f_R(r) = 2h^2 r e^{-h^2 r^2} \mathbb{1}_{r>0}$ où $h > 0$ est un paramètre réel fixé. Rappelons que $\mathbb{1}_{r>0}$ égal 1 si $r > 0$ et 0 si $r \leq 0$.

1. Calculer la distance moyenne entre la position d'alunissage de la fusée et la position prévue.
2. Calculer la variance de cette distance.
3. Donner une interprétation du paramètre h .
4. Que doit valoir h pour que dans plus de 90 % des alunissages, la fusée se pose à moins de 200 mètres de la position prévue ?

1. La distance moyenne entre la position d'alunissage et la position prévue par les astronautes est $\mathbb{E}(R) = \int_{\mathbb{R}} r f_R(r) dr = 2h^2 \int_0^{+\infty} r^2 e^{-h^2 r^2} dr = \left[-r e^{-h^2 r^2} \right]_0^{+\infty} + \int_0^{+\infty} e^{-h^2 r^2} dr$ par une intégration par partie. Le terme entre crochets est nul. Calculons le second terme. Le changement de variables $x = hr$ conduit à $\int_0^{+\infty} e^{-h^2 r^2} dr = \frac{I}{2h}$, où $I = \int_{\mathbb{R}} e^{-x^2} dx$.



Le calcul de $I = \int_{\mathbb{R}} e^{-x^2} dx$ est classique. Il consiste à calculer I^2 en passant en coordonnées polaires. Nous avons $I^2 = \int_{\mathbb{R}} e^{-x^2} dx \times \int_{\mathbb{R}} e^{-y^2} dy = \int_{\mathbb{R}^2} e^{-(x^2+y^2)} dx dy$. Posons maintenant $x = r \cos \theta$, $y = r \sin \theta$ avec $r > 0$ et $\theta \in]0; 2\pi[$, alors $I^2 = \int_0^{2\pi} \int_0^{+\infty} e^{-r^2} r dr d\theta = \left(\int_0^{2\pi} d\theta \right) \times \left(\int_0^{+\infty} r e^{-r^2} dr \right) = 2\pi \times \frac{1}{2}$. D'où $I = \sqrt{\pi}$.



Fiche 45



Fiche 65

Finalement, $\mathbb{E}(R) = \frac{I}{2h} = \frac{\sqrt{\pi}}{2h}$.

2. Calculons à présent $\text{Var}(R) = \mathbb{E}(R^2) - \mathbb{E}(R)^2$. Nous avons besoin de $\mathbb{E}(R^2) = \int_{\mathbb{R}} r^2 f_R(r) dr = 2h^2 \int_0^{+\infty} r^3 e^{-h^2 r^2} dr = \left[-r^2 e^{-h^2 r^2}\right]_0^{+\infty} + 2 \int_0^{+\infty} r e^{-h^2 r^2} dr$ par intégration par parties. Le terme entre crochets est nul, et le terme intégral $2 \int_0^{+\infty} r e^{-h^2 r^2} dr$ est égal à $\frac{2}{2h^2} = \frac{1}{h^2}$. Finalement, $\mathbb{E}(R^2) = \frac{1}{h^2}$ et $\text{Var}(R) = \frac{1}{h^2} \left(1 - \frac{\pi}{4}\right)$.

3. Finalement, la distance à la cible a une moyenne et une variance qui s'expriment respectivement en $\frac{1}{h}$ et en $\frac{1}{h^2}$. Nous en déduisons que plus h est grand, plus $\mathbb{E}(R)$ et $\text{Var}(R)$ sont petits. Autrement dit, plus h est grand, plus les astronautes sont habiles pour poser la fusée sur la position prévue (car $\mathbb{E}(R)$ est petit), et ils sont précis (car $\text{Var}(R)$ est petit).

4. Nous voulons que la fusée se pose à moins de 200 mètres de la position prévue, et ce dans plus de 90 % des cas. Ceci s'écrit $\mathbb{P}(R \leq 200) \geq 90\%$. Or, $\mathbb{P}(R \leq 200) = F_R(200) = \int_{-\infty}^{200} f_R(r) dr = \left[-e^{-h^2 r^2}\right]_0^{200} = 1 - e^{-200^2 h^2}$. La résolution de $1 - e^{-200^2 h^2} \geq 90\%$ conduit à $h \geq \frac{\sqrt{-\ln 10}}{200}$, soit encore $h \geq 0,0076$. Il est logique de trouver h supérieur à une certaine valeur, car nous avons vu que plus h est grand, plus les astronautes sont habiles et précis pour poser la fusée.

Pour s'entraîner (solutions p. 544)

78.1 Chez une personne en bonne santé, le taux de diabète à jeun doit se situer entre 0,8 et 1,2 g/L et varier peu. Deux personnes attendent chez le médecin. Soient X_1 et X_2 leurs taux de diabète respectifs, à jeun. Ces deux personnes ont un taux de diabète moyen de 0,9 g/L : $\mathbb{E}(X_1) = \mathbb{E}(X_2) = 0,9$ g/L. Mais leurs écarts-types sont différents : $\sqrt{\text{Var}(X_1)} = 0,1$ g/L et $\sqrt{\text{Var}(X_2)} = 1$ g/L. Des deux personnes, laquelle doit absolument se faire soigner pour son diabète ?

78.2 Chaque année, une action rapporte un taux de 50 %, 20 % ou 10 % avec probabilités respectives $p_{50} = 0,1$, $p_{20} = 0,7$ et $p_{10} = 0,2$. Soit X le taux de l'action. Calculer l'espérance et la variance de X .

78.3 Soit X une variable aléatoire réelle. Montrer le Théorème de Huygens :

$$\text{Var}(X) = \mathbb{E}\left((X - \mathbb{E}(X))^2\right) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2.$$

Pourquoi la variance est-elle positive ?

78.4 Soit a un réel et X une variable aléatoire constante égale à a , c'est-à-dire $\mathbb{P}(X = a) = 1$. Montrer que $\text{Var}(X) = 0$.

78.5 Soient a et b deux réels, et X une variable aléatoire réelle. Montrer que $\text{Var}(aX + b) = a^2 \text{Var}(X)$. Nous notons $\mu = \mathbb{E}(X)$ et $\sigma^2 = \text{Var}(X)$. Montrer que $\mathbb{E}\left(\frac{X - \mu}{\sigma}\right) = 0$ et $\text{Var}\left(\frac{X - \mu}{\sigma}\right) = 1$.

78.6 Montrer que si X et Y sont deux variables aléatoires réelles de même loi, et admettant une densité, alors elles ont les mêmes moments.

Paramètres de position Intervalles de probabilité

Outre l'espérance et la variance, nous présentons ici d'autres grandeurs permettant de décrire le comportement d'une variable aléatoire réelle X . Ces grandeurs ne suffisent pas à caractériser la loi des variables aléatoires, mais elles en donnent une description synthétique.

Fiches 77
et 78

1. Mode

Définition 79.1

Le **mode** d'une variable aléatoire réelle X est la valeur $x_m \in X(\Omega)$ telle que $P_{x_m} = \sup \{p_x; x \in X(\Omega)\}$ si X est discrète ou $f_x(x_m) = \sup \{f_x(x); x \in X(\Omega)\}$ si X est à densité.

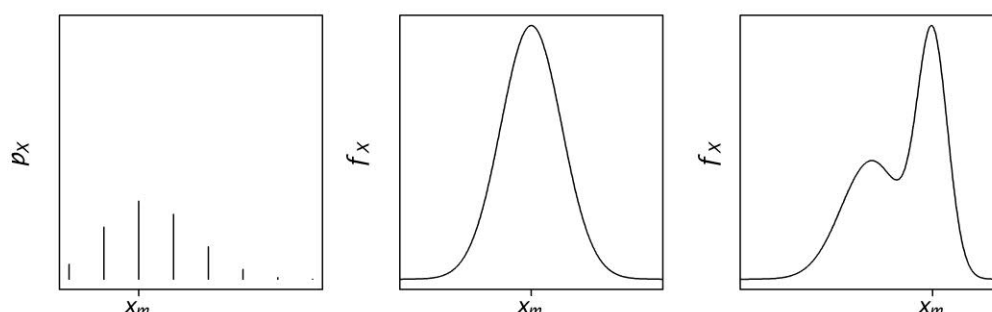


Figure 79.1 – Mode x_m : cas d'une variable discrète (à gauche), et à densité (au milieu et à droite).



- Le mode n'est pas toujours défini. Par exemple, si X a pour densité $f_X(x) = \mathbb{1}_{x \in [0;1]}$, c'est-à-dire $f_X(x) = 1$ si $x \in [0;1]$ et 0 sinon, le maximum de f_X est atteint sur $[0;1]$.
- S'il existe plusieurs maxima locaux pour la loi de la variable aléatoire X (cf. Figure 79.1, graphe de droite), X est **multimodale** ou **plurimodale** (par exemple, s'il y a deux maxima locaux, X est dite **bimodale**). S'il n'existe qu'un maximum, X est **unimodale**.

2. Médiane

Définition 79.2

La **médiane** d'une variable aléatoire réelle X est la valeur $m \in \mathbb{R}$ qui sépare l'ensemble $X(\Omega)$ en deux parties de même probabilité $1/2$, c'est-à-dire la valeur m telle que :

$$\mathbb{P}(X \leq m) = \mathbb{P}(X > m) = \frac{1}{2}.$$

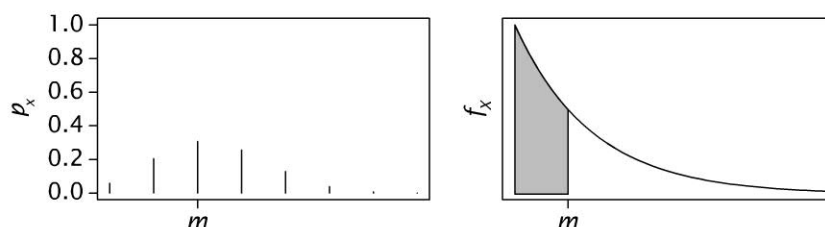


Figure 79.2 – Médiane : cas d’une variable discrète (à gauche), et à densité (à droite). Dans le graphique de droite, les aires hachurée et non hachurée sous la courbe valent toutes deux $\frac{1}{2}$.



Point méthode

La médiane m est donc solution de l’équation $F_X(x) = \frac{1}{2}$, où F_X désigne la fonction de répartition (définie à la Fiche 76) de la variable aléatoire X .

- Dans le cas d’une variable aléatoire discrète,
 - soit l’équation $F_X(x) = \frac{1}{2}$ n’a pas de solution. La médiane est alors définie par la valeur m telle que $\mathbb{P}(X < m) < \frac{1}{2}$ et $\mathbb{P}(X \geq m) > \frac{1}{2}$ (graphique de gauche dans la Figure 79.3).
 - soit elle admet un ensemble de solutions de la forme $[m_1; m_2[$ (graphique du milieu dans la Figure 79.3).
- Dans le cas d’une variable aléatoire à densité, si F_X est une fonction continue et strictement croissante, alors la solution de $F_X(x) = \frac{1}{2}$ est unique : m est cette unique solution (graphique de droite dans la Figure 79.3)

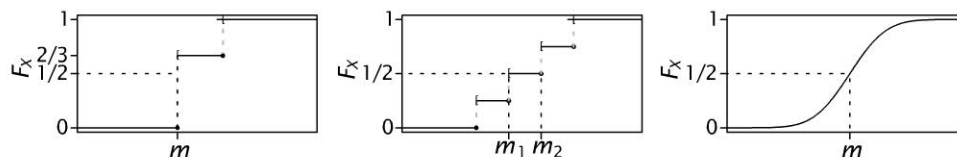


Figure 79.3 – Médiane : cas d’une variable discrète (à gauche et au milieu), et à densité (à droite).

3. Quantiles

Définition 79.3

Soit $p \in [0; 1]$. Le **quantile d’ordre p** d’une variable aléatoire réelle X est défini par la valeur x_p telle que $\mathbb{P}(X \leq x_p) = p$. Autrement dit, x_p est solution de l’équation $F_X(x) = p$, d’inconnue x .



- La médiane est donc le quantile d’ordre 0,5.
- Dans le cas d’une variable aléatoire à densité, si F_X est continue et strictement croissante, alors pour tout $p \in]0; 1[$, la solution de $F_X(x) = p$ existe toujours, et est unique : $x_p = F_X^{-1}(p)$, où F_X^{-1} est la fonction réciproque de F_X (voir le Théorème 36.7 dans la partie Analyse).

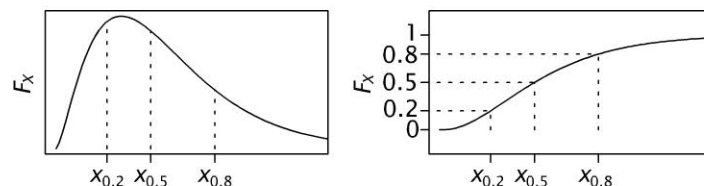


Figure 79.4 – Quantiles d’ordre 0,2 ; 0,5 et 0,8 pour une variable aléatoire X à densité : schémas avec la densité de X (à gauche), et la fonction de répartition de X (à droite).

4. Intervalle de probabilité

Définition 79.4

Soient $\alpha \in [0; 1]$ et $a, b \in \mathbb{R} \cup \{\pm\infty\}$. L'intervalle $[a; b]$ est un **intervalle de probabilité** de niveau $1 - \alpha$ pour la variable aléatoire réelle X si $\mathbb{P}(X \in [a; b]) = \mathbb{P}(X \leq b) - \mathbb{P}(X < a) = 1 - \alpha$.

Exemple

1. Dans la Figure 79.4, l'intervalle $[x_{0,2}; x_{0,8}]$ est un intervalle de probabilité de X de niveau 60 %.
2. L'intervalle de probabilité n'est pas déterminé de manière unique. Par exemple, la Figure 79.5 donne trois intervalles de probabilité de niveau 90 % pour une même variable aléatoire X de densité f_X .

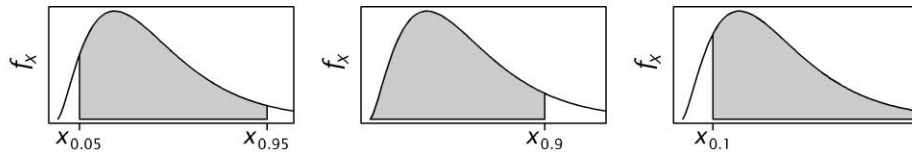


Figure 79.5 – Intervalles de probabilité de niveau 90 % pour la variable aléatoire X . Nous les qualifions de bilatéral (graphe de gauche), unilatéral gauche (graphe du milieu) ou unilatéral droite (graphe de droite).

Application La température en Lorraine

Soit X une variable aléatoire réelle de densité définie pour tout $x \in \mathbb{R}$ par $f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$, dont nous admettons dans cette Fiche que $-1,96$ et $1,96$ sont les quantiles d'ordre 2,5 % et 97,5 %. Soit T la température moyenne (en degrés Celsius) en juillet en Lorraine. Nous supposons que $T = 19 + 3X$. Donner un intervalle de probabilité bilatéral de niveau 95 % de la température moyenne en juillet en Lorraine. Donner la loi, le mode et la médiane de T .

1. Écrivons la donnée de l'énoncé : 95 % = $\mathbb{P}(-1,96 \leq X \leq 1,96)$.

Maintenant nous cherchons deux valeurs t_1 et t_2 telles que 95 % = $\mathbb{P}(t_1 \leq T \leq t_2)$. Nous avons $T = 19 + 3X$ donc 95 % = $\mathbb{P}(t_1 \leq 19 + 3X \leq t_2) = \mathbb{P}\left(\frac{t_1 - 19}{3} \leq X \leq \frac{t_2 - 19}{3}\right)$. Par identification, nous obtenons $\frac{t_1 - 19}{3} = -1,96$ et $\frac{t_2 - 19}{3} = 1,96$, ce qui donne $t_1 = 13,12$ et $t_2 = 24,79$.

2. Calculons la fonction de répartition F_T de $T = 19 + 3X$. Nous avons $F_T(t) = \mathbb{P}(T \leq t) = \mathbb{P}(19 + 3X \leq t) = \mathbb{P}\left(X \leq \frac{t - 19}{3}\right) = F_X\left(\frac{t - 19}{3}\right)$.



L'équation $F_T(t) = F_X\left(\frac{t - 19}{3}\right)$ permet aussi de trouver les quantiles de $T = 19 + 3X$ à partir de ceux de X . Par exemple, $F_X(-1,96) = 2,5$ % d'après l'énoncé. Si nous cherchons par exemple t_1 tel que $F_T(t_1) = 2,5$ % alors il suffit d'écrire $F_T(t_1) = F_X\left(\frac{t_1 - 19}{3}\right) = 2,5$ % et identifier $\frac{t_1 - 19}{3}$ avec $-1,96$.

Par dérivation de F_T nous obtenons la densité $f_T(t) = F'_T(t) = \frac{1}{3}F'_X\left(\frac{t-19}{3}\right) = \frac{1}{3\sqrt{2\pi}}e^{-(t-19)^2/18}$.

3. Le maximum de f_T est atteint en $t = 19$. Le mode de T est donc égal à 19. La température la plus fréquente est donc 19 degrés. Par symétrie de la densité de T , nous remarquons que $\int_{-\infty}^{19} \frac{1}{3\sqrt{2\pi}}e^{-(t-19)^2/18}dt = \int_{19}^{+\infty} \frac{1}{3\sqrt{2\pi}}e^{-(t-19)^2/18}dt$. Or, la somme de ces deux termes est égale à l'aire sous la courbe de la densité f_T , c'est-à-dire 1. Nous obtenons donc que ces deux termes valent tous les deux $\frac{1}{2}$. Nous remarquons qu'ils sont égaux respectivement à $\mathbb{P}(T \leq 19)$ et $\mathbb{P}(T \geq 19)$. La température médiane en Lorraine est donc de 19 degrés Celsius.

Pour s'entraîner (solutions p. 544)

79.1 Soit X le résultat d'un lancer de dé équilibré. La variable aléatoire X a-t-elle un mode ?

79.2 Soit X la durée de vie d'une batterie, de fonction de répartition $F_X(t) = (1 - e^{-\lambda t}) \mathbb{1}_{t \geq 0}$ avec $\lambda > 0$ un paramètre fixé.

1. Quelle est la durée de vie médiane de la batterie ?
2. Le constructeur dit que dans 90 % des cas, sa batterie dure au moins 10 unités de temps. Cette valeur est un quantile : de quel ordre est-il ? En déduire λ .
3. Donner un intervalle de probabilité bilatéral à 95 % de la durée de vie de la batterie.

79.3 Soit X la variable aléatoire de densité

$$f_X(x) = \frac{1}{\sqrt{2\pi}} \exp(-x^2/2), \quad \text{avec } x \in \mathbb{R}.$$

Nous disons que X suit la loi $\mathcal{N}(0; 1)$

1. Calculer le mode de X .
2. Calculer la médiane de X .
3. Nous savons que 1,960 est le quantile d'ordre 0,975 de X . Quel est le quantile d'ordre 0,025 de X ?
4. Donner un intervalle de probabilité bilatéral de niveau 95 % pour X .

 Fiche 83

Fonction d'une variable aléatoire réelle – Calcul de loi

Il arrive souvent que nous ne connaissions pas la loi de la variable aléatoire réelle qui nous intéresse, disons Y , mais que nous connaissions la loi d'une autre variable aléatoire réelle X , telle que $Y = h(X)$ pour une certaine fonction $h : \mathbb{R} \rightarrow \mathbb{R}$.

Exemple

1. C'est le cas dans l'exercice détaillé de la Fiche 79, où la température $T = 19 + 3X$ est exprimée en fonction d'une autre variable aléatoire X de loi connue.
2. Supposons qu'une machine produit des résistances dont la valeur en ohms est une variable aléatoire R de loi connue. Une intensité i fixée traverse la résistance, nous voulons savoir la loi de la tension U définie par $U = R \times i$.

Nous présentons différentes méthodes pour obtenir la loi de $Y = h(X)$, à partir de celle de X , qu'elle soit discrète ou à densité.

1. Cas où $Y = h(X)$ est discrète

Chercher la loi de Y revient à identifier $Y(\Omega)$ et déterminer $\mathbb{P}(Y = y)$ pour tout $y \in Y(\Omega)$.

Il faut alors identifier, pour tout $y \in Y(\Omega)$, l'événement $\{h(X) = y\}$ équivalent à l'événement $\{Y = y\}$. Ensuite, nous calculons $\mathbb{P}(Y = y) = \mathbb{P}(h(X) = y)$.

Exemple

Soit X le diamètre intérieur d'une rondelle fabriquée en usine. Nous supposons connue la loi de X , et nous notons f_X sa densité. L'usage de ces rondelles nécessite que le diamètre varie entre deux bornes de tolérance 1,2 mm et 1,4 mm. Si le diamètre ne vérifie pas cette condition, la rondelle est défectueuse. Soit $Y = \mathbb{1}_{X \in [1,2; 1,4]}$ la variable aléatoire qui vaut 1 si la rondelle est conforme aux bornes de tolérance, et 0 sinon. Y est une variable aléatoire discrète :

$$Y(\Omega) = \{0; 1\}, \mathbb{P}(Y = 1) = \mathbb{P}(X \in [1, 2; 1, 4]) = \int_{1,2}^{1,4} f_X(x) dx, \text{ et } \mathbb{P}(Y = 0) = \mathbb{P}(X < 1, 2 \text{ ou } X >$$

$$1, 4) = 1 - \int_{1,2}^{1,4} f_X(x) dx.$$

Nous verrons à la Fiche 81 que la loi de Y est une loi de Bernoulli de paramètre $p = \int_{1,2}^{1,4} f_X(x) dx$.

2. Cas où $Y = h(X)$ est continue

Pour définir la loi de Y , nous pouvons par exemple chercher sa fonction de répartition F_Y définie par $F_Y(y) = \mathbb{P}(Y \leq y)$, pour tout $y \in \mathbb{R}$.

Il faut alors chercher l'ensemble $A \subset \mathbb{R}$ tel que l'événement $\{Y \leq y\}$ soit équivalent à l'événement $\{X \in A\}$. Ensuite, nous calculons $\mathbb{P}(Y \leq y) = \mathbb{P}(X \in A)$.

Exemple

Soit un signal radio Y modélisé par une sinusoïde : $Y = \cos(X)$ où X est la phase du signal. Cette phase est supposée aléatoire de densité f_X définie sur $[-\pi; \pi]$. Nous voyons que

$$Y(\Omega) \subset [-1; 1], \text{ et pour tout } y \in [-1; 1], \mathbb{P}(Y \leq y) = \mathbb{P}(X \in [\arccos y; \pi] \cup [-\pi; -\arccos y]) = \int_{\arccos y}^{\pi} f_X(x) dx + \int_{-\pi}^{-\arccos y} f_X(x) dx.$$

► Cas où $Y = h(X)$ avec h strictement monotone

Supposons que la variable aléatoire X a une densité connue, notée f_X . Notons Y la variable aléatoire égale à $h(X)$. Si h est strictement monotone sur $I = X(\Omega)$, nous rappelons que h est alors bijective de I sur $h(I)$.

Cas où h est strictement croissante sur I

Dans ce cas, pour tout $y \in h(I)$, la fonction de répartition de Y s'obtient ainsi :

$$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(h(X) \leq y) = \mathbb{P}(X \leq h^{-1}(y)) = \int_{-\infty}^{h^{-1}(y)} f_X(x) dx.$$

Exemple

Reprenons l'exemple de $U = R \times i$ avec R de densité f_R supposée connue, définie sur $\mathbb{R}^+$. L'intensité i est fixée. La loi de U est donnée par $F_U(u) = \mathbb{P}(U \leq u) = \mathbb{P}\left(R \leq \frac{u}{i}\right) = \int_0^{\frac{u}{i}} f_R(r) dr$, si $u \geq 0$ et $F_U(u) = 0$ si $u < 0$. La densité de U est alors $f_U(u) = F'_U(u) = \frac{1}{i} f_R\left(\frac{u}{i}\right) \mathbb{1}_{u \geq 0}$.

Cas où h est strictement décroissante sur I

Pour tout $y \in h(I)$, la fonction de répartition de Y s'obtient ainsi :

$$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(h(X) \leq y) = \mathbb{P}(X \geq h^{-1}(y)) = \int_{h^{-1}(y)}^{+\infty} f_X(x) dx = 1 - \int_{-\infty}^{h^{-1}(y)} f_X(x) dx.$$

Dans les deux cas (h strictement croissante ou décroissante), la densité de Y s'obtient par dérivation de la fonction F_Y en tout point où F_Y est dérivable, c'est-à-dire en tout point $y \in \mathbb{R}$ où h^{-1} est définie et dérivable :

$$f_Y(y) = f_X(h^{-1}(y)) \left| \frac{dh^{-1}(y)}{dy} \right|.$$

3. Moments d'une fonction d'une variable aléatoire réelle

Soit $Y = h(X)$, où X est une variable aléatoire réelle de loi connue, et Y une variable aléatoire de loi inconnue.

Le calcul de $\mathbb{E}(h(X))$ et de $\text{Var}(h(X))$ a été défini dans les Fiches 77 et 78.

 Fiche 36

 Fiches 77 et 78

Considérons une machine produisant des résistances dont la valeur ohmique X (en ohms) est une variable aléatoire de densité f_X supposée connue et définie sur $\mathbb{R}^+$. Un appareil mesure la valeur des résistances produites, mais ne fonctionne plus pour des résistances supérieures à 1000 ohms. Nous notons Y la valeur donnée par l'appareil. Quelle est la loi de la variable aléatoire Y ?

Nous avons $Y = \min(X; 1000)$. Calculons la fonction de répartition de Y .

- Si $y < 1000$, l'événement $\{Y \leq y\}$ est équivalent à l'événement $\{X \leq y\}$.

$$\text{Donc } \mathbb{P}(Y \leq y) = \mathbb{P}(X \leq y) = \int_0^y f_X(x) dx.$$

- Si $y = 1000$, l'événement $\{Y = 1000\}$ est équivalent à l'événement $\{X \geq 1000\}$ puisque l'appareil sature à 1000 ohms.

$$\text{Donc } \mathbb{P}(Y = 1000) = \mathbb{P}(X \geq 1000) = \int_{1000}^{+\infty} f_X(x) dx.$$

- Enfin, si $y \geq 1000$, $\mathbb{P}(Y \leq y) = 1$ car y est nécessairement inférieur à 1000.

Nous avons ainsi obtenu la loi de Y via sa fonction de répartition. Notons ici que Y est à la fois discrète (car elle prend la valeur 1000 avec une probabilité non nulle) et continue (sur l'intervalle $[0; 1000[$). Nous disons que Y est une variable aléatoire « mixte ».

Pour s'entraîner (solutions p. 545)

80.1 Soit la variable aléatoire X de loi définie par $X(\Omega) = \{3; 5; 8\}$ avec les probabilités $\mathbb{P}(X = 3) = 0,3$, $\mathbb{P}(X = 5) = 0,3$ et $\mathbb{P}(X = 8) = 0,4$. Soit $Y = a + bX$. Quelle est la loi de Y ?

80.2 Nous reprenons l'énoncé de l'exercice détaillé de la Fiche 78, où R désigne l'écart (en mètres) entre la position d'alunissage de la fusée et la position prévue par les astronautes. Soit $S = \pi R^2$ la surface du disque dans lequel la fusée s'est posée.

1. Donner la fonction de répartition de S .
2. Donner la densité de S .

80.3 Soit X la longueur d'un tissu enroulé sur un rouleau et destiné à être coupé en bouts de trois mètres. La densité de X est connue, notée f_X et définie sur $\mathbb{R}^+$. Soit Y le nombre de bouts de trois mètres que peut donner le tissu. Quelle est la loi de Y ?

80.4 Le jeu de la roulette au casino possède 37 cases numérotées de 0 à 36. Suppo-

sons que cette roulette n'est pas truquée : tous les numéros sont équiprobables.

1. Un joueur mise 1 euro sur un nombre pair : si un nombre pair sort, il rempoche sa mise et gagne x euro(s) supplémentaire(s). Sinon, il perd sa mise. Au casino, le nombre 0 n'est pas considéré comme nombre pair. Notons X son gain. Quelle est sa loi ? Que doit valoir x pour que le joueur ait en moyenne un gain positif ?

2. Le joueur mise maintenant 1 euro sur le numéro 0. Si le 0 sort, il rempoche sa mise et gagne y euros supplémentaires. Sinon, il perd sa mise. Notons Y son gain. Quelle est sa loi ? Que doit valoir y pour que le joueur ait en moyenne un gain positif ?

3. Le joueur mise ensuite 1 euro sur un nombre pair, et 2 euros sur le 0. Si un nombre pair sort, il rempoche sa mise de 1 euro et gagne z euros supplémentaires. Si le 0 sort, il rempoche sa mise de 2 euros et gagne $10z$ euros supplémentaires. Sinon, il perd ses deux mises. Notons Z son gain. Quelle est sa loi ?

Que doit valoir z pour que le joueur ait en moyenne un gain positif ?

80.5 Soient $\lambda > 0$, et X une variable aléatoire de densité $f_X(x) = \lambda \exp(-\lambda x) \mathbb{1}_{x \geq 0}$.

1. Calculer la fonction de répartition F_X de X .
2. Calculer l'inverse G de F_X , définie par la fonction $G : [0; 1[\rightarrow \mathbb{R}^+$ telle que $F_X(G(u)) = u$ pour tout $u \in [0; 1[$.
3. Montrer que si U est une variable aléatoire de densité $f_U(u) = \mathbb{1}_{0 \leq u < 1}$, alors $G(U)$ a la même loi que X .

80.6 

Nous reprenons l'application de la Fiche 77, avec X le nombre pains demandés, y le nombre de pains fabriqués par le boulanger, f_X la densité de X , G le gain du boulanger, et les paramètres de bénéfice et de déficit β et δ . Calculer la loi de la variable aléatoire G en fonction de F_X , β , δ et y où F_X désigne la fonction de répartition de X . *Indication : fixer une valeur de y , et utiliser la formule des*

probabilités totales, avec la partition d'événements $\{X \leq y\}$ et $\{X > y\}$.

80.7 Un missile est tiré du point origine (cf. Figure 80.1). La portée R du tir est liée à l'angle de tir α (en radians), à la vitesse v (en m/s) du missile à la sortie du canon selon la relation

$$R = \frac{v^2}{g} \sin(2\alpha)$$

où $g = 9,81 \text{ m/s}^2$. Du fait des possibles erreurs de positionnement du canon, l'angle α varie aléatoirement entre $[\frac{\pi}{4} - \frac{\pi}{360}; \frac{\pi}{4} + \frac{\pi}{360}]$ selon la loi

$$f_\alpha(t) = \frac{180}{\pi} \mathbb{1}_{t \in [\frac{\pi}{4} - \frac{\pi}{360}; \frac{\pi}{4} + \frac{\pi}{360}]},$$

où nous rappelons que $\mathbb{1}_{t \in A}$ est égal à 1 si $t \in A$ et 0 sinon. Quelle est la loi de R ? *Indication : pour tout $r \in R(\Omega)$, identifier α_0 et α_1 dans $\alpha(\Omega)$ tels que les événements $\{R \leq r\}$ et $\{\alpha \in [\frac{\pi}{4} - \frac{\pi}{360}; \alpha_0] \cup [\alpha_1; \frac{\pi}{4} + \frac{\pi}{360}]\}$ soient égaux.*

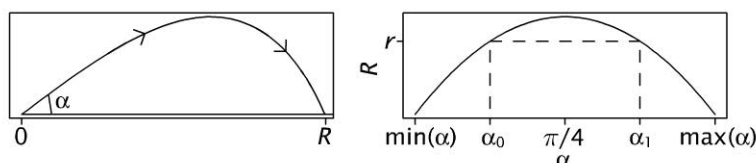


Figure 80.1 – Trajectoire d'un missile (à gauche) et portée en fonction de l'angle de tir (à droite).

Lois discrètes usuelles à valeurs dans un ensemble fini

Nous présentons ci-dessous les lois classiques de variables aléatoires réelles discrètes prenant un ensemble fini de valeurs.

1. Loi de Dirac

Soit $a \in \mathbb{R}$. La **loi de Dirac** en a est la loi d'une **variable aléatoire** X **constante** égale à a :

$$X(\Omega) = \{a\}, \quad \mathbb{P}(X = a) = 1.$$

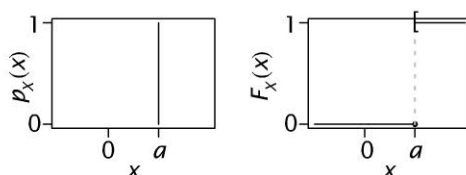


Figure 81.1 – Loi de Dirac, avec ici $a > 0$: diagramme en bâtons à gauche, fonction de répartition à droite.

2. Loi uniforme discrète $\mathcal{U}\{1; \dots; n\}$

Soit $n \in \mathbb{N}^*$. La **loi uniforme discrète** sur $\{1; \dots; n\}$, notée $\mathcal{U}\{1; \dots; n\}$, est la loi d'une variable aléatoire X prenant chacune des valeurs de l'ensemble $\{1; \dots; n\}$ avec la même probabilité :

$$X(\Omega) = \{1; \dots; n\}, \quad \forall k \in \{1; \dots; n\}, \quad \mathbb{P}(X = k) = \frac{1}{n}.$$

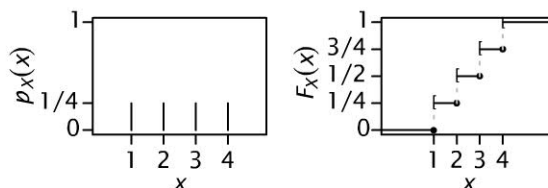


Figure 81.2 – Loi $\mathcal{U}\{1; 2; 3; 4\}$: diagramme en bâtons à gauche, fonction de répartition à droite.

3. Loi de Bernoulli $\mathcal{B}(p)$

Soit $p \in [0; 1]$. La **loi de Bernoulli** de paramètre p , notée $\mathcal{B}(p)$, est la loi d'une variable aléatoire X prenant la valeur 1 (ou succès) avec probabilité p , et la valeur 0 (ou échec) avec probabilité $1 - p$:

$$X(\Omega) = \{0; 1\}, \quad \mathbb{P}(X = 0) = 1 - p, \quad \mathbb{P}(X = 1) = p.$$

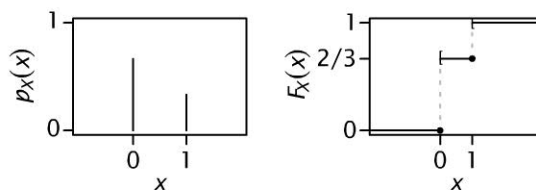


Figure 81.3 – Loi $\mathcal{B}\left(\frac{1}{3}\right)$: diagramme en bâtons à gauche, fonction de répartition à droite.

4. Loi binomiale $\mathcal{B}(n; p)$

Soient $n \in \mathbb{N}^*$ et $p \in [0; 1]$. Considérons une urne contenant des boules de type A en proportion p et des boules de type B en proportion $1-p$, et imaginons que nous effectuons n **tirages avec remise** dans cette urne. Le nombre X de boules de type A ainsi tirées est une variable aléatoire de **loi binomiale** de paramètres $(n; p)$. Elle est notée $\mathcal{B}(n; p)$:

$$X(\Omega) = \{0; \dots; n\}, \quad \forall k \in X(\Omega), \quad \mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

Le coefficient $\binom{n}{k}$ a été défini à la Fiche 71. Notons que l'événement $\{X = k\}$ signifie que k boules de type A ont été obtenues parmi les n boules tirées avec remise.



Autrement dit, la loi binomiale $\mathcal{B}(n; p)$ est la loi du nombre de succès parmi n expériences aléatoires **indépendantes, et réalisées dans les mêmes conditions**.

Ici, le succès serait de tirer une boule de type A.

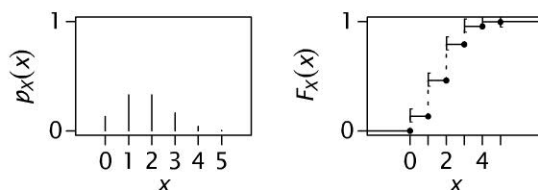


Figure 81.4– Loi $\mathcal{B}\left(5; \frac{1}{3}\right)$: diagramme en bâtons à gauche, fonction de répartition à droite.



Soit $p \in [0; 1]$. La loi de Bernoulli de paramètre p est la loi binomiale de paramètres $(1; p)$.

5. Loi hypergéométrique $\mathcal{H}(N; p; n)$

Soit E une urne contenant N boules, dont une proportion p est de type A, et une proportion $1-p$ est de type B. Nous **tirons sans remise** $n \leq N$ boules. Le nombre X de boules de type A ainsi tirées est une variable aléatoire de **loi hypergéométrique**, notée $\mathcal{H}(N; p; n)$:

$$X(\Omega) = \{\sup(0; n - (1-p)N); \dots; \inf(pN; n)\},$$

$$\forall k \in X(\Omega), \quad \mathbb{P}(X = k) = \frac{\binom{pN}{k} \times \binom{(1-p)N}{n-k}}{\binom{N}{n}}.$$

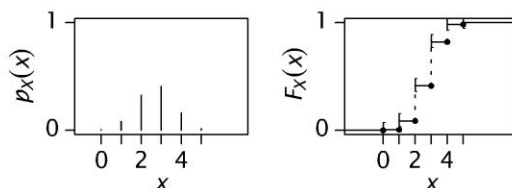


Figure 81.5– Loi $\mathcal{H}\left(13; \frac{7}{13}; 5\right)$: diagramme en bâtons à gauche, fonction de répartition à droite.

Soit un boulevard bordé de bancs publics distants de 1 mètre. Un ivrogne se déplace de bancs en bancs, par pas de 1 mètre, vers la droite ou vers la gauche, sans direction privilégiée. Les bancs sont numérotés $-k; -(k-1); \dots; -1; 0; 1; \dots; k-1; k$. À l'instant initial, l'ivrogne est au niveau du banc 0. Quelle est la position moyenne de l'ivrogne après qu'il a fait n pas ? Quelle est la variance de cette position ?

1. Commençons par noter P_n la variable aléatoire indiquant le numéro du banc où se trouve l'ivrogne après qu'il a fait n pas. La position moyenne de l'ivrogne après avoir fait n pas est $\mathbb{E}(P_n)$. Pour connaître cette position moyenne, une première idée est de chercher la loi de P_n .

2. Par exemple, $P_1(\Omega) = \{-1; 1\}$ avec $\mathbb{P}(P_1 = 1) = \mathbb{P}(P_1 = -1) = \frac{1}{2}$. De même, $P_2(\Omega) = \{-2; 0; 2\}$, avec $\mathbb{P}(P_2 = 2) = \mathbb{P}(P_2 = -2) = \frac{1}{4}$ et $\mathbb{P}(P_2 = 0) = \frac{1}{2}$. Nous pouvons continuer par récurrence.

3. Mais ici, il sera plus simple de pratiquer autrement. En effet, notons D_n (respectivement G_n) le nombre de pas vers la droite (resp. vers la gauche). Alors la position de l'ivrogne après n pas s'exprime par

$$P_n = D_n - G_n,$$

et d'autre part, l'ivrogne faisant n pas au total, nous avons

$$D_n + G_n = n,$$

d'où

$$P_n = 2D_n - n.$$

Examinons D_n . Si un « succès » est de faire un pas à droite, et un « échec » un pas à gauche, alors D_n est le nombre de « succès » parmi n expériences aléatoires indépendantes, et réalisées dans les mêmes conditions. La probabilité de « succès » est ici de $\frac{1}{2}$. Nous en déduisons que D_n suit la loi binomiale de paramètres n et $\frac{1}{2}$.

4. Nous pouvons donc conclure, par la linéarité de l'espérance : $\mathbb{E}(P_n) = 2\mathbb{E}(D_n) - n = 2 \times \frac{n}{2} - n = 0$, en utilisant la formule de l'espérance d'une loi binomiale. L'ivrogne, en moyenne, sera sur le banc initial : ce qui n'est pas étonnant, car ses pas à gauche ou à droite ont les mêmes probabilités $\frac{1}{2}$.

5. Enfin, $\text{Var}(P_n) = \text{Var}(2D_n - n) = 4 \times \text{Var}(D_n) = 4 \times \frac{n}{4} = n$. C'est une grande variance : la position de l'ivrogne varie beaucoup.



Fiche 95

Pour s'entraîner (solutions p. 545)

81.1 Soient 12 forages, suffisamment distants pour être indépendants. Chaque forage a une probabilité $\frac{1}{4}$ de contenir du pétrole. Soit X le nombre de forages sur lesquels nous trouvons du pétrole.

1. Quelle est la loi de la variable aléatoire X ?
2. Combien de forages, en moyenne, donneront du pétrole ?

81.2 Dans une fête foraine, un gérant de pêche aux canards a 100 lots, dont 25 jouets et 75 sachets de bonbons. Un enfant joue 10 fois. À chaque fois qu'il joue, il garde son lot pour lui.

1. Quelle est la probabilité qu'il pêche 5 jouets et 5 sachets de bonbons ?
2. Quelle est la probabilité qu'il pêche 1 jouet et 9 sachets de bonbons ?

81.3 Soit X une variable aléatoire de Bernoulli, de paramètre $p \in [0; 1]$.

1. Calculer la moyenne et la variance de X .
2. Pour quelle valeur de p la variance est-elle maximale ?

81.4 Soit un dé non truqué. Nous lançons ce dé une fois et notons X le résultat du dé.

1. Quelle est la loi de la variable aléatoire X ?
2. Calculer son espérance et sa variance.

81.5 Soit un dé non truqué. Nous lançons ce dé n fois et notons X le nombre de 6 obtenus.

1. Quelle est la loi de la variable aléatoire X ?
2. Quelles sont son espérance et sa variance ?

81.6 Soit X une variable aléatoire de loi de Dirac en $a \in \mathbb{R}$. Calculer la moyenne et la variance de X .

Lois discrètes usuelles à valeurs dans un ensemble dénombrable

Nous présentons ci-dessous les lois classiques de variables aléatoires réelles discrètes prenant un ensemble dénombrable de valeurs.

1. Loi de Poisson $\mathcal{P}(\lambda)$

Soit $\lambda \in \mathbb{R}_+^*$. La **loi de Poisson** de paramètre λ est la loi d'une variable aléatoire X comptant le nombre d'occurrences d'un événement sur une unité de temps (ou de longueur, de surface, de volume, etc.). Nous avons :

$$X(\Omega) = \mathbb{N}, \quad \forall k \in \mathbb{N}, \quad \mathbb{P}(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}.$$



Le paramètre λ est le nombre moyen d'occurrences par unité de temps (ou de longueur, de surface, de volume, etc.) : c'est-à-dire que nous avons $\mathbb{E}(X) = \lambda$.

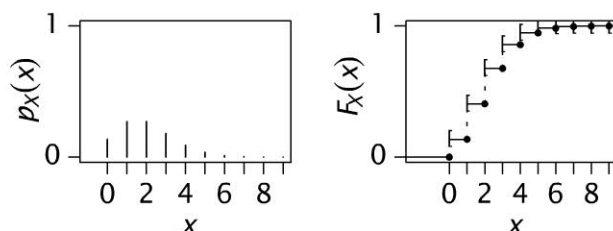


Figure 82.1 – Loi $\mathcal{P}(2)$: diagramme en bâtons à gauche, fonction de répartition à droite.

2. Loi géométrique $\mathcal{G}(p)$

Soit $p \in]0; 1[$. Considérons une urne contenant des boules de type A en proportion p et des boules de type B en proportion $1 - p$, et imaginons que nous effectuons des **tirages avec remise** dans cette urne jusqu'à tirer une boule de type A . Le nombre X de tirages nécessaires pour obtenir une boule de type A est une variable aléatoire de **loi géométrique** de paramètre p . Elle est notée $\mathcal{G}(p)$. Nous avons :

$$X(\Omega) = \mathbb{N}^*, \quad \forall k \in \mathbb{N}^*, \quad \mathbb{P}(X = k) = p(1 - p)^{k-1}.$$



La loi géométrique $\mathcal{G}(p)$ est la loi du nombre d'essais nécessaires pour obtenir un succès lors d'une série d'expériences aléatoires indépendantes, dont la probabilité de succès est p .

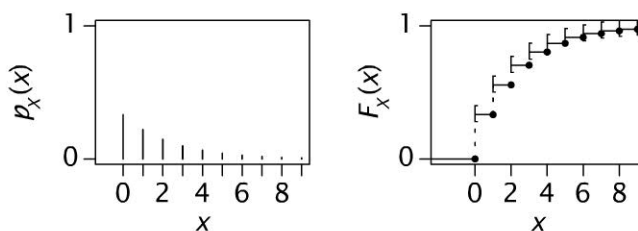


Figure 82.2 – Loi $\mathcal{G}(\frac{1}{3})$: diagramme en bâtons à gauche, fonction de répartition à droite.

3. Loi binomiale négative

Soient $r \in \mathbb{N}^*$ et $p \in]0; 1[$. Considérons à nouveau une urne contenant des boules de type A en proportion p et des boules de type B en proportion $1 - p$, et imaginons que nous effectuons des **tirages avec remise** dans cette urne jusqu'à tirer r boules de type A . Le nombre X de tirages nécessaires pour obtenir r boules de type A est une variable aléatoire de **loi binomiale négative** de paramètres $(r; p)$. Nous avons :

$$X(\Omega) = r; r + 1; \dots, \quad \forall k \in X(\Omega), \quad \mathbb{P}(X = k) = \binom{k-1}{r-1} p^r (1-p)^{k-r},$$

où le coefficient $\binom{k-1}{r-1}$ a été défini à la Fiche 71.



La loi binomiale négative de paramètres $(r; p)$ est la loi du nombre d'essais nécessaires pour obtenir r succès lors d'une série d'expériences aléatoires indépendantes, dont la probabilité de succès est p .



La loi géométrique de paramètre p est une loi binomiale négative de paramètres $(1; p)$.

Application L'ours

Mon grand-père est parti au Canada dans le but d'observer un ours. Chaque jour, la probabilité d'observer l'ours est de $\frac{1}{5}$ (l'apparition de l'ours un jour donné ne dépend pas des autres jours). Mon grand-père se poste chaque jour jusqu'à ce qu'il voit l'ours. Combien de jours doit-il s'attendre à patienter ? Quel est l'écart-type de ce nombre de jours ?

1. La variable aléatoire considérée, notée X , est le nombre de jours que mon grand-père doit patienter pour voir l'ours. Chaque jour, l'expérience aléatoire « voir l'ours » est réalisée avec probabilité $p = \frac{1}{5}$. Nous avons une suite d'expériences aléatoires indépendantes, chacune ayant une probabilité p de succès. X est le nombre d'essais nécessaires pour obtenir un succès, donc $X \sim \mathcal{G}(p)$.

2. Le nombre de jours durant lesquels mon grand-père doit s'attendre à patienter est $\mathbb{E}(X)$: c'est le nombre moyen de jours nécessaires pour voir apparaître l'ours.

3. Calculons l'espérance d'une loi géométrique : $\mathbb{E}(X) = \sum_{k=1}^{+\infty} k \mathbb{P}(X = k) = \sum_{k=1}^{+\infty} k p (1-p)^{k-1}$.



Le calcul de cette série repose sur une astuce de calcul. Nous savons que la fonction $f(x) = \frac{1}{1-x}$ définie pour $x \neq 1$ s'écrit sous forme de série, si la condition $|x| < 1$ est vérifiée :

$$\frac{1}{1-x} = \sum_{k=0}^{+\infty} x^k.$$

Les résultats de la Fiche 50 permettent de dire que la somme de cette série de fonctions est dérivable sur $] -1; 1[$ et que :

$$\frac{1}{(1-x)^2} = \sum_{k=1}^{+\infty} k x^{k-1}. \quad (82.1)$$

Appliquons cette formule avec $x = 1 - p$, nous obtenons si $p \in]0; 1[$:

$$\sum_{k=1}^{+\infty} k(1-p)^{k-1} = \frac{1}{(1-(1-p))^2} = \frac{1}{p^2}.$$

Finalement, $\mathbb{E}(X) = p \times \frac{1}{p^2} = \frac{1}{p}$: mon grand-père devra attendre en moyenne 5 jours avant de voir l'ours.

4. À présent, calculons $\text{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$. Le calcul direct de $\mathbb{E}(X^2) = \sum_{k=1}^{+\infty} k^2 p(1-p)^{k-1}$ est difficile. Nous proposons de calculer $\mathbb{E}(X(X-1)) = \sum_{k=1}^{+\infty} k(k-1)p(1-p)^{k-1}$.



De même que précédemment, nous pouvons dériver sur $] - 1; 1[$ la fonction définie à l'Équation (82.1) :

$$\frac{2}{(1-x)^3} = \sum_{k=2}^{+\infty} k(k-1)x^{k-2}.$$

Posons $x = 1 - p$, nous obtenons si $p \in]0; 1[$:

$$\sum_{k=2}^{+\infty} k(k-1)(1-p)^{k-2} = \frac{2}{p^3}.$$

Finalement, $\mathbb{E}(X(X-1)) = p(1-p) \times \frac{2}{p^3} = \frac{2(1-p)}{p^2}$. Or $\mathbb{E}(X(X-1)) = \mathbb{E}(X^2) - \mathbb{E}(X)$, nous en déduisons $\mathbb{E}(X^2)$, et finalement $\text{Var}(X) = \frac{1-p}{p^2}$. La variance du temps nécessaire pour voir l'ours est $\sqrt{20}$ jours, c'est-à-dire environ 4,5 jours.

Pour s'entraîner (solutions p. 546)

82.1 Soit X une variable aléatoire suivant une loi de Poisson, de paramètre $\lambda > 0$. Calculer la moyenne et la variance de X .

82.2 Notons X_t le nombre d'éruptions du volcan Etna sur un intervalle de temps de longueur $t > 0$ (en années). Supposons que X_t suit la loi de Poisson de paramètre λt avec $\lambda > 0$.

1. Quel est le nombre moyen d'éruptions sur l'intervalle de temps $[0; t]$?
2. Admettons que l'Etna a en moyenne 100 éruptions par siècle. Que choisir pour le paramètre λ ?
3. Quelle est la probabilité d'observer au moins une éruption de l'Etna sur les cinq ans à venir ? Même question sur le siècle à venir ?

4. Quel est le nombre d'éruptions le plus probable pouvant survenir pendant une année ?

82.3 Notons X le nombre de défauts le long d'un câble de caténaire destiné à alimenter des trains électriques. Dans une région donnée, le nombre moyen de défauts par mètre de câble est donné par $\mu = 4 \times 10^{-5} \text{ m}^{-1}$. Nous supposons que la loi de X est une loi de Poisson de paramètre $\mu \ell$ où ℓ est la longueur du câble de caténaire considéré (en mètres).

1. Sur un caténaire de 200 km, combien de défauts devons-nous attendre ?
2. Quel est le nombre de défauts le plus probable sur une telle longueur ?

3. Quelle est la probabilité de n'avoir aucun défaut sur le câble ?

82.4 Dans un certain pays se trouvent des mines, dont 10 % contiennent de l'or, 5 % contiennent des diamants, et 2 % contiennent à la fois de l'or et des diamants.

1. Combien de mines doit parcourir un chercheur d'or et de diamants avant de découvrir de l'or et/ou des diamants ?

2. Même question pour un chercheur uniquement d'or.

82.5 Une société vend des appareils ménagers. Elle sait que chaque appareil a une probabilité p de tomber en panne chaque année (nous supposons que cette probabilité reste

constante au cours des années). Cette société propose une garantie d'une durée t années. Quelle doit-être la valeur de t pour que la société n'ait pas à réparer plus de 25 % des appareils vendus ?

82.6 Une poule voudrait trois poussins. Lorsqu'elle pond un œuf, celui-ci a une probabilité $p = \frac{1}{4}$ d'éclore et de donner un poussin vivant.

1. Combien d'œufs, en moyenne, la poule doit-elle pondre pour que trois de ces œufs éclosent ?

2. Quelle est la probabilité que la poule « dusse » pondre 10 œufs pour voir enfin éclore son troisième œuf ?

De nombreux phénomènes aléatoires suivent approximativement une loi normale (par exemple, la concentration en nitrate dans l'eau d'une nappe phréatique, l'erreur de mesure d'un appareil, etc.). La raison de cette importance sera vue à la Fiche 93.

1. Loi normale centrée et réduite $\mathcal{N}(0; 1)$

La variable aléatoire réelle X suit la **loi normale centrée et réduite** ou **loi normale standard**, notée $\mathcal{N}(0; 1)$, si X a pour densité :

$$\forall x \in \mathbb{R}, \quad f_X(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right).$$

Nous notons $X \sim \mathcal{N}(0; 1)$. La variable aléatoire X vérifie $\mathbb{E}(X) = 0$ et $\text{Var}(X) = 1$.



- La densité de la loi normale centrée et réduite est symétrique, centrée sur 0.
- La fonction de répartition de la loi normale centrée et réduite, souvent notée Φ , est donnée par :

$$\forall x \in \mathbb{R}, \quad \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp\left(-\frac{t^2}{2}\right) dt.$$

Théorème 83.1

- La fonction de répartition Φ de la loi normale standard vérifie : $\forall x > 0, \quad \Phi(-x) = 1 - \Phi(x)$.
- Les quantiles de la loi normale centrée et réduite sont définis par inversion de la fonction Φ . Pour tout $p \in]0; 1[$, le quantile d'ordre p est l'unique $z \in \mathbb{R}$ tel que $\Phi(z) = p$.

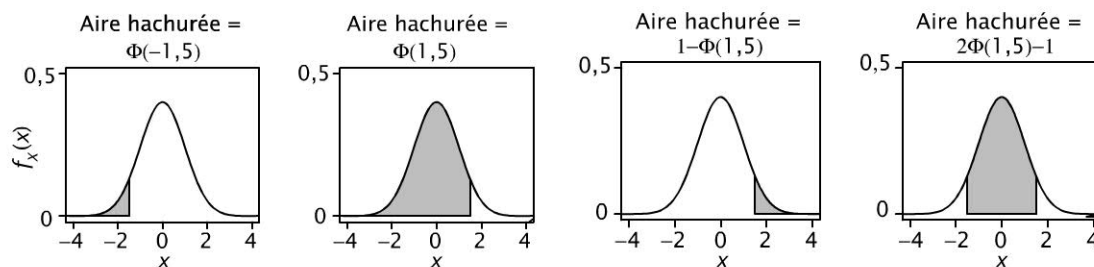


Figure 83.1 – Illustration de la propriété $\Phi(x) = 1 - \Phi(-x)$ avec ici $x = -1,5$.

2. Loi normale $\mathcal{N}(\mu; \sigma)$

Soient $\mu \in \mathbb{R}$ et $\sigma > 0$. La variable aléatoire réelle X suit la **loi normale** de paramètres μ et σ , notée $\mathcal{N}(\mu; \sigma)$, si X a pour densité :

$$\forall x \in \mathbb{R}, \quad f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right).$$

Nous notons $X \sim \mathcal{N}(\mu; \sigma)$. La variable aléatoire X vérifie $\mathbb{E}(X) = \mu$ et $\text{Var}(X) = \sigma^2$.



- La densité de la loi normale $\mathcal{N}(\mu; \sigma)$ est symétrique, centrée sur μ .
- Le paramètre σ^2 représente la **variance** : plus σ est grand, plus la densité de la loi normale est **dispersée**, comme le montre la Figure 83.3.
- Dans la littérature, la loi normale de moyenne μ et de variance σ^2 est parfois notée $\mathcal{N}(\mu; \sigma^2)$.

Théorème 83.2

Soient X une variable aléatoire de loi $\mathcal{N}(\mu; \sigma)$, et a, b deux réels. Alors $Y = aX + b$ suit la loi normale $\mathcal{N}(a\mu + b; |a|\sigma)$. En particulier, nous avons

$$\frac{X - \mu}{\sigma} \sim \mathcal{N}(0; 1).$$

Si $X \sim \mathcal{N}(\mu; \sigma)$, la fonction de répartition de X est donc définie par :

$$\forall x \in \mathbb{R}, \quad F_X(x) = \mathbb{P}(X \leq x) = \mathbb{P}\left(\frac{X - \mu}{\sigma} \leq \frac{x - \mu}{\sigma}\right) = \Phi\left(\frac{x - \mu}{\sigma}\right),$$

et pour tout $p \in]0; 1[$, le quantile d'ordre p de X est l'unique $z \in \mathbb{R}$ tel que $\Phi\left(\frac{z - \mu}{\sigma}\right) = p$.

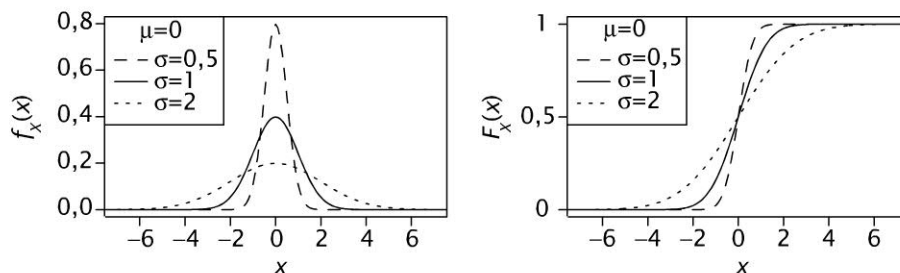


Figure 83.2 – Loi $\mathcal{N}(\mu; \sigma)$: densité à gauche, fonction de répartition à droite pour diverses valeurs de μ . La loi $\mathcal{N}(0; 1)$ est en trait plein.

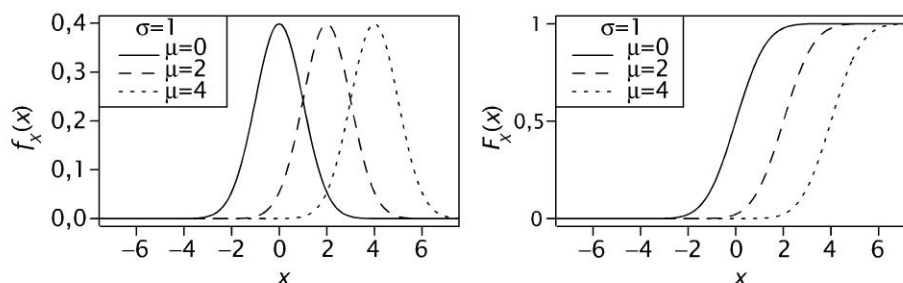


Figure 83.3 – Loi $\mathcal{N}(\mu; \sigma)$: densité à gauche, fonction de répartition à droite pour diverses valeurs de σ . La loi $\mathcal{N}(0; 1)$ est en trait plein.

Dans une entreprise, le salaire X (en euros) suit une loi normale $\mathcal{N}(1000; 50)$.

1. Quelle est la probabilité de gagner moins de 950 euros par mois ? Plus de 1050 euros par mois ?
2. Donner un intervalle de probabilité à 95 % du salaire des employés, de la forme $[1000 \pm x]$.
3. Un quota de 5 % des plus faibles salaires de cette entreprise bénéficie d'une aide au logement. Quel est le salaire à ne pas dépasser pour avoir cette aide ?
4. Dans une autre entreprise, le salaire Y est réparti selon une loi $\mathcal{N}(1000; \sigma)$ et 90 % des employés de cette entreprise gagnent entre 800 et 1200 euros par mois. Que vaut σ ?



Il faut prendre l'habitude, lorsque $X \sim \mathcal{N}(\mu; \sigma)$, de se ramener à une loi normale centrée et réduite par la transformation $\frac{X - \mu}{\sigma}$.

Notons Z une variable aléatoire de loi $\mathcal{N}(0; 1)$.

1. Soit $X \sim \mathcal{N}(1000; 50)$. Nous cherchons $\mathbb{P}(X \leq 950)$ et $\mathbb{P}(X \geq 1050)$.

$$\begin{aligned}\mathbb{P}(X \leq 950) &= \mathbb{P}\left(\frac{X - 1000}{50} \leq \frac{950 - 1000}{50}\right) \\ &= \mathbb{P}(Z \leq -1) = \Phi(-1) = 1 - \Phi(1) \\ &= 1 - 0,8413 = 0,1587 \text{ par lecture dans la table de la loi normale centrée et réduite.}\end{aligned}$$

De même, $\mathbb{P}(X \geq 1050) = \mathbb{P}\left(\frac{X - 1000}{50} \geq \frac{1050 - 1000}{50}\right) = \mathbb{P}(Z \geq 1) = \mathbb{P}(Z \leq -1)$ **par symétrie de la loi normale centrée et réduite**. Donc, $\mathbb{P}(X \geq 1050) = 0,1587$.

2. D'après la table de la loi normale centrée et réduite, nous savons que

$$0,95 = \mathbb{P}(-1,960 \leq Z \leq 1,960).$$

Donc, nous avons

$$\begin{aligned}0,95 &= \mathbb{P}(-1,960 \leq Z \leq 1,960) \\ &= \mathbb{P}\left(-1,960 \leq \frac{X - 1000}{50} \leq 1,960\right) \\ &= \mathbb{P}(1000 - 1,960 \times 50 \leq X \leq 1000 + 1,960 \times 50) \\ &= \mathbb{P}(902 \leq X \leq 1098).\end{aligned}$$

L'intervalle de probabilité à 95 % du salaire X dans cette entreprise est $[902; 1098]$ euros.

3. Cherchons le quantile d'ordre 5 % pour la variable aléatoire X , c'est-à-dire le réel q tel que $\mathbb{P}(X \leq q) = 0,05 \Leftrightarrow \mathbb{P}(X \leq -q) = 0,95$. La table de la loi normale centrée et réduite nous donne $\mathbb{P}(Z \leq -1,645) = 0,05$. Ainsi,

$$\begin{aligned}5 \% &= \mathbb{P}(Z \leq -1,645) \\ &= \mathbb{P}\left(\frac{X - 1000}{50} \leq -1,645\right) \\ &= \mathbb{P}(X \leq 1000 - 1,645 \times 50).\end{aligned}$$

Le quantile q cherché est donc $q = 1000 - 1,645 \times 50 = 917,75$ euros. C'est le salaire à ne pas dépasser pour bénéficier de l'aide au logement.

4. L'énoncé indique que $0,90 = \mathbb{P}(800 \leq Y \leq 1200)$. Nous avons donc

$$\begin{aligned} 0,90 &= \mathbb{P}(800 \leq Y \leq 1200) \\ &= \mathbb{P}\left(\frac{800 - 1000}{\sigma} \leq \frac{Y - 1000}{\sigma} \leq \frac{1200 - 1000}{\sigma}\right) \\ &= \mathbb{P}\left(-\frac{200}{\sigma} \leq Z \leq \frac{200}{\sigma}\right). \end{aligned}$$

Or, nous savons, d'après la table de la loi normale centrée et réduite que $90 \% = \mathbb{P}(-1,645 \leq Z \leq 1,645)$. Nous en déduisons que $\frac{200}{\sigma} = 1,645$ et donc que $\sigma = \frac{200}{1,645} = 121,5805$ euros. Dans cette seconde entreprise, les salaires sont plus dispersés que dans la précédente.

Pour s'entraîner (solutions p. 547)

83.1 Soit Z une variable aléatoire de loi $\mathcal{N}(0; 1)$.

1. Déterminer z_1, z_2, z_3 tels que

$$\mathbb{P}(Z \leq z_1) = 97,5 \%,$$

$$\mathbb{P}(Z \leq z_2) = 2,5 \%$$

$$\text{et } \mathbb{P}(-z_3 \leq Z \leq z_3) = 90 \%.$$

2. Calculer $\mathbb{P}(Z \leq 1,64)$ et $\mathbb{P}(Z \leq -1,64)$.

3. Calculer $\mathbb{P}(|Z| < 1)$, $\mathbb{P}(-2 < Z < 1)$ et $\mathbb{P}(|Z - 1| > 0,5)$.

4. Calculer $\mathbb{E}(Z)$ et $\mathbb{V}\text{ar}(Z)$.

83.2 Soit X une variable aléatoire de loi $\mathcal{N}(\mu; \sigma)$ avec $\mu = 1$ et $\sigma = 2$.

1. Déterminer x_1, x_2, x_3 tels que

$$\mathbb{P}(X \leq x_1) = 97,5 \%,$$

$$\mathbb{P}(X \leq x_2) = 2,5 \%$$

$$\text{et } \mathbb{P}(-x_3 \leq X - \mu \leq x_3) = 90 \%.$$

2. Que devient l'intervalle $[-x_3; x_3]$ si σ devient petit ?

3. Calculer $\mathbb{P}(X \leq 4,28)$, $\mathbb{P}(2 < X < 3)$, $\mathbb{P}(X > 0)$ et $\mathbb{P}(|X - 3| > 1)$.

4. Calculer $\mathbb{E}(X)$ et $\mathbb{V}\text{ar}(X)$.

83.3 Mon entreprise propose deux types d'actions. Un investissement de 100 euros dans la première action me rapportera $100(1 + X_1)$ euros, avec $X_1 \sim \mathcal{N}(0, 2; 0, 5)$; tandis qu'il me rapportera $100(1 + X_2)$ euros dans la seconde, avec $X_2 \sim \mathcal{N}(0, 8; 1)$. En considérant l'intervalle de probabilité bilatéral de niveau 95 % de ce que je peux recevoir dans les deux cas, et si je ne veux pas endetter ma famille, dans quelle action vaut-il mieux que j'investisse 100 euros ?

Lois continues usuelles

Outre la loi normale, il existe quantité de lois continues. Rappelons que $\mathbb{1}_{x \in A} = 1$ si $x \in A$, et 0 sinon.

1. Loi uniforme $\mathcal{U}[a; b]$

Soient $a < b$ deux réels. La variable aléatoire X à valeurs dans $[a; b]$, suit la **loi uniforme** sur $[a; b]$, notée $\mathcal{U}[a; b]$, si X a pour densité :

$$\forall x \in \mathbb{R}, \quad f_X(x) = \frac{1}{b-a} \mathbb{1}_{x \in [a; b]}(x).$$

La fonction de répartition de X est définie sur $\mathbb{R}$ par : $F_X(x) = \frac{x-a}{b-a} \mathbb{1}_{x \in [a; b]} + \mathbb{1}_{x \geq b}$.

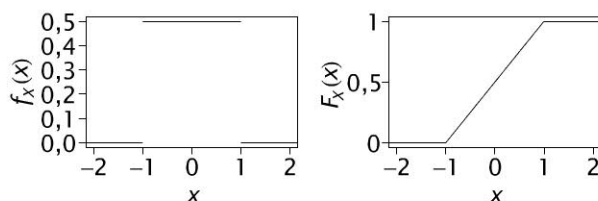


Figure 84.1 – Loi $\mathcal{U}[-1; 1]$: densité à gauche, fonction de répartition à droite.



La loi uniforme $\mathcal{U}[a; b]$ est la loi de la position d'un point jeté « au hasard » dans le segment $[a; b]$ (c'est-à-dire sans privilégier aucune partie du segment $[a; b]$).

2. Loi exponentielle $\mathcal{E}(\lambda)$

Soit $\lambda > 0$. La variable aléatoire X à valeurs positives, suit la **loi exponentielle** de paramètre λ , notée $\mathcal{E}(\lambda)$, si X a pour densité :

$$\forall x \in \mathbb{R}, \quad f_X(x) = \lambda \exp(-\lambda x) \mathbb{1}_{x \geq 0}.$$

La fonction de répartition de X est définie sur $\mathbb{R}$ par : $F_X(x) = [1 - \exp(-\lambda x)] \mathbb{1}_{x \geq 0}$.

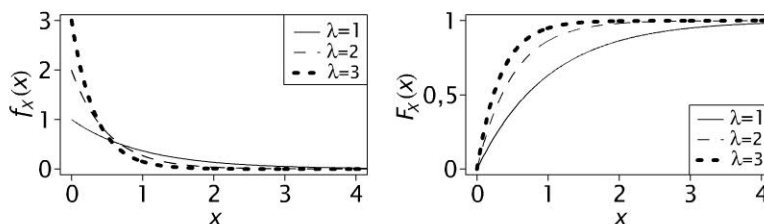


Figure 84.2 – Loi $\mathcal{E}(\lambda)$: densité à gauche, fonction de répartition à droite, pour divers choix de λ .



La loi exponentielle $\mathcal{E}(\lambda)$ est la loi d'une durée de vie d'un appareil qui ne vieillit pas.

3. Loi de Weibull $W(a; b)$

Soient $a > 0, b > 0$. La variable aléatoire X à valeurs positives, suit la **loi de Weibull** de paramètres a, b , notée $W(a; b)$, si X a pour densité :

$$\forall x \in \mathbb{R}, \quad f_X(x) = \frac{a}{b^a} x^{a-1} e^{-(x/b)^a} \mathbb{1}_{x \geq 0}.$$

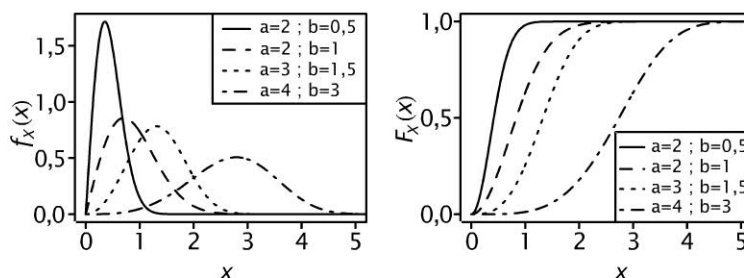


Figure 84.3 – Loi $W(a; b)$: densité à gauche, fonction de répartition à droite, pour divers choix de a, b .



- La loi exponentielle $\mathcal{E}(\lambda)$ est un cas particulier de la loi de Weibull : c'est la loi $W(1; 1/\lambda)$.
- La loi de Weibull représente une durée de vie d'un objet. Si $a > 1$, l'objet a un taux de panne qui augmente : l'objet s'use. Si $a < 1$, l'objet a un taux de panne qui diminue : l'objet se « rode ». Si $a = 1$, l'objet a un taux de panne constant : l'objet ne vieillit pas.

4. Loi gamma $\Gamma(a; b)$

Soient $a > 0, b > 0$. La variable aléatoire X à valeurs positives, suit la **loi gamma** de paramètres a et b , notée $\Gamma(a; b)$, si X a pour densité :

$$\forall x \in \mathbb{R}, \quad f_X(x) = \frac{1}{b^a \Gamma(a)} x^{a-1} e^{-x/b} \mathbb{1}_{x \geq 0},$$

où $\Gamma(a) = \int_0^{+\infty} t^{a-1} e^{-t} dt$ est la fonction Gamma d'Euler.

Fiche 53

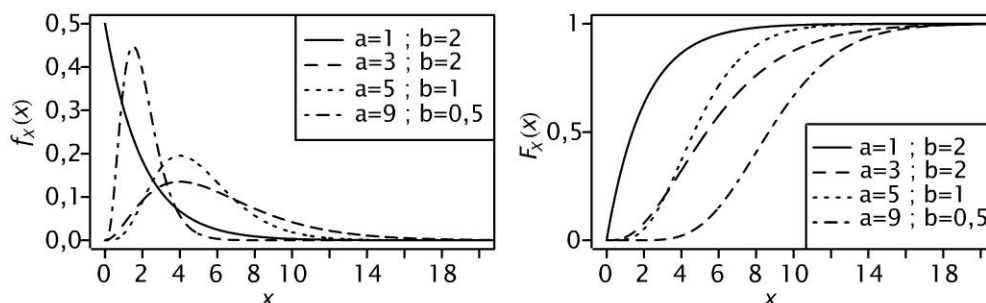


Figure 84.4 – Loi $\Gamma(a; b)$: densité à gauche, fonction de répartition à droite pour divers choix de a et b .



La loi exponentielle $\mathcal{E}(\lambda)$ est un cas particulier de la loi gamma : c'est la loi $\Gamma(1; 1/\lambda)$.

5. Loi beta $Beta(a; b)$

Soient $a > 0, b > 0$. La variable aléatoire X à valeurs dans $[0; 1]$ suit la **loi beta** de paramètres a et b , notée $Beta(a; b)$, si X a pour densité :

$$\forall x \in \mathbb{R}, \quad f_X(x) = \frac{1}{B(a; b)} x^{a-1} (1-x)^{b-1} \mathbb{1}_{x \in [0; 1]},$$

où $B(a; b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} = \int_0^1 x^{a-1} (1-x)^{b-1} dx$ est la fonction Beta d'Euler.

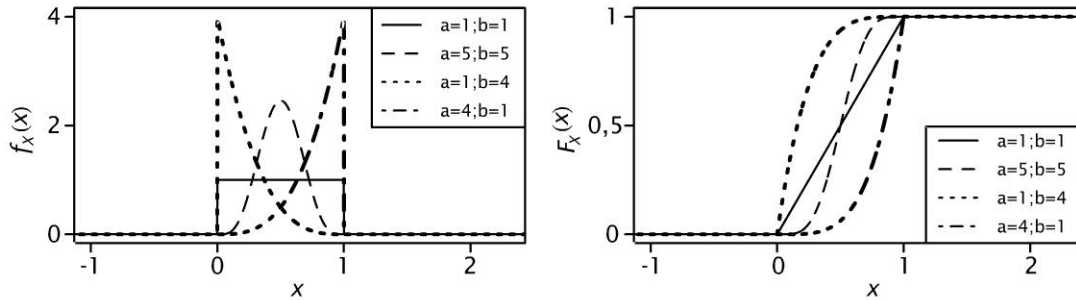


Figure 84.5 – Loi $Beta(a; b)$: densité à gauche, fonction de répartition à droite pour divers choix de a et b .



La loi uniforme $\mathcal{U}[0; 1]$ est un cas particulier de la loi beta : c'est la loi $Beta(1; 1)$.

Application La durée de vie d'une batterie

Soit X une variable aléatoire représentant la durée de vie d'une batterie (en jours). Nous supposons que X suit une loi $\mathcal{E}(\lambda)$ avec $\lambda = \frac{1}{900}$ jours⁻¹. Calculer la durée de vie moyenne de cette batterie. Quelle est la probabilité qu'elle dure plus de 3 ans ? Si la batterie a déjà duré plus de $s > 0$ jours, quelle est la probabilité qu'elle dure encore plus de $t > 0$ jours ? Commenter ce résultat. À partir de quel instant $t_{90\%}$ la batterie a-t-elle 90 % de « chance » d'être en panne ?

1. La durée moyenne de la batterie est $\mathbb{E}(X) = \int_0^{+\infty} \lambda x e^{-\lambda x} dx = \frac{1}{\lambda} = 900$ jours, obtenu par intégration par parties.

2. La probabilité que la batterie dure plus de 3 ans, soit 1095 jours si nous négligeons les années bissextiles, est $\mathbb{P}(X \geq 1095) = \int_{1095}^{+\infty} \lambda x e^{-\lambda x} dx = e^{-1095\lambda} = 0,296$. Remarquons que c'est aussi $1 - F_X(1095)$.

3. Si la batterie a déjà duré plus de $s > 0$ jours, la probabilité qu'elle dure encore plus de $t > 0$ jours est $\mathbb{P}(X \geq t + s | X \geq s) = \frac{\mathbb{P}(X \geq t + s, X \geq s)}{\mathbb{P}(X \geq s)} = \frac{\mathbb{P}(X \geq t + s)}{\mathbb{P}(X \geq s)}$ car l'événement $\{X \geq t + s, X \geq s\}$ est égal à l'événement $\{X \geq t + s\}$. Finalement, $\mathbb{P}(X \geq t + s | X \geq s) = \frac{1 - F_X(t + s)}{1 - F_X(s)} = \frac{e^{-\lambda(t+s)}}{e^{-\lambda s}} = e^{-\lambda t}$. Nous voyons que cette probabilité ne dépend pas de s , le

 Fiche 45

temps qu'a déjà duré la batterie. C'est pourquoi la loi exponentielle est dite loi des durées de vie **sans vieillissement** ou **sans mémoire**.

4. La valeur $t_{90\%}$ est le quantile d'ordre 90 % : $\mathbb{P}(X \leq t_{90\%}) = 0,9$. Il faut donc résoudre $1 - e^{-\lambda t_{90\%}} = 0,9$, donc $t_{90\%} = -\frac{\ln 0,1}{\lambda} = 2072,3$ jours soit à peu près 5 ans et demi.

Pour s'entraîner (solutions p. 547)

84.1 Calculer l'espérance et la variance de chacune des lois présentées dans cette Fiche. Nous indiquons que les fonctions Gamma et Beta d'Euler vérifient $\Gamma(x+1) = x\Gamma(x)$ et $B(x+1; y) = \frac{x}{x+y}B(x; y)$, pour tous réels $x > 0, y > 0$.

84.2 Considérons les instants d'éruptions $T_1 < T_2 < \dots$ d'un volcan à partir de T_0 fixé au 01/01/1900. Soient $S_1, S_2, \dots$ les durées (en jours) entre les éruptions : S_1 est l'instant de la première éruption (depuis T_0), S_2 est la durée entre la deuxième et la première éruption, S_i est la durée entre la $i^{\text{ème}}$ et la $i-1^{\text{ème}}$ éruption.

1. Quelle est la relation entre T_k et S_i pour $i = 1; \dots; k$?

2. Supposons que S_i suit une loi exponentielle de paramètre λ avec $\lambda = \frac{1}{365}$ jours⁻¹, pour tout $i \geq 1$. Nous admettons que T_k suit une loi gamma de paramètres k et $1/\lambda$, pour tout $k \geq 1$. Calculer $\mathbb{E}(S_i)$ et $\mathbb{E}(T_k)$.

84.3

www

Soit X l'instant d'arrivée du courrier à ma boîte aux lettres.

Je sais que le facteur passe au hasard entre 9 h et 10 h.

1. Quelle est la loi de X ?
2. Quelle est la fonction de répartition de X ?
3. Quelle est la probabilité que mon courrier ne soit pas encore arrivé à 9^h45 ?
4. Quelle est la probabilité que mon courrier ne soit pas encore arrivé à 9^h30, mais qu'il soit arrivé avant à 9^h45 ?

84.4

www

Montrer que la loi $\mathcal{E}(\lambda)$ est un cas particulier de la loi gamma et de la loi de Weibull.

84.5

www

Montrer que la loi $\mathcal{U}([0; 1])$ est un cas particulier de la loi beta.

Lois continues usuelles (suite)

Les quantiles des lois présentées dans cette Fiche sont donnés dans les tables à la fin de ce livre.

1. Loi du Khi-deux $\chi^2(n)$

Soit $n \in \mathbb{N}^*$. La variable aléatoire X , à valeurs positives, suit la **loi du Khi-deux** à n degrés de liberté, notée $\chi^2(n)$, si X a pour densité :

$$\forall x \in \mathbb{R}, \quad f_X(x) = \frac{1}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)} x^{\frac{n}{2}-1} \exp\left(-\frac{x}{2}\right) \mathbb{1}_{x \geq 0},$$

où $\Gamma(r) = \int_0^{+\infty} t^{r-1} e^{-t} dt$ est la fonction Gamma d'Euler définie à la Fiche 53.

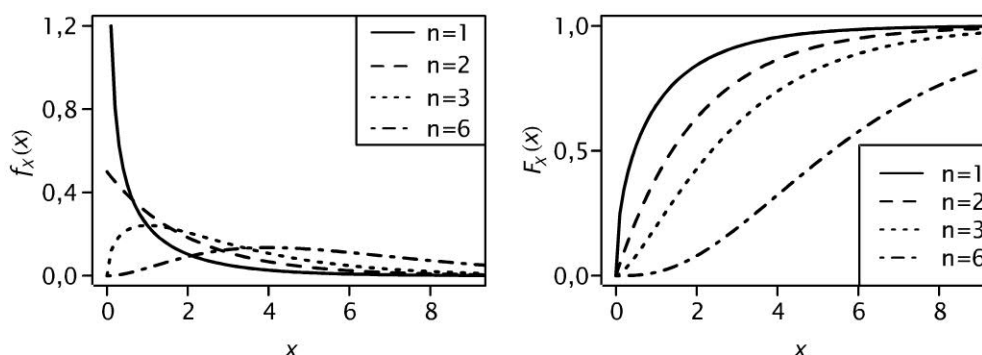


Figure 85.1 – Loi $\chi^2(n)$ pour différents choix de n : densité à gauche, fonction de répartition à droite.



La loi du Khi-deux à n degrés de liberté est un cas particulier de la loi gamma : c'est la loi gamma de paramètres $\frac{n}{2}$ et 2.

2. Loi de Student $t(n)$

Soit $n \in \mathbb{N}^*$. La variable aléatoire X , à valeurs réelles, suit la loi de **Student** à n degrés de liberté, notée $t(n)$, si X a pour densité :

$$\forall x \in \mathbb{R}, \quad f_X(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi} \Gamma\left(\frac{n}{2}\right)} \times \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}.$$

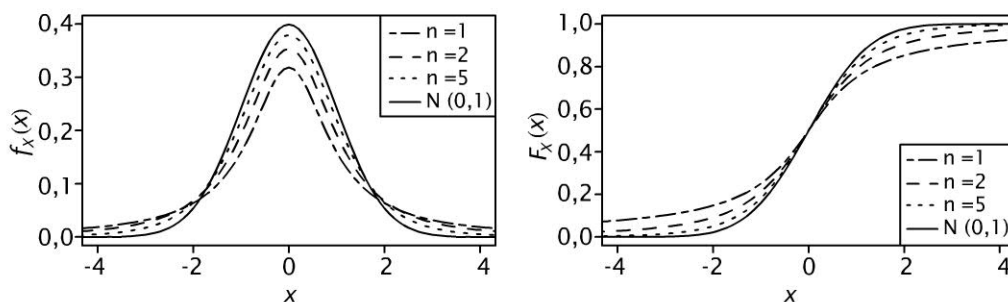


Figure 85.2 – Loi $t(n)$ pour différents choix de n : densité à gauche, fonction de répartition à droite.



Lorsque n tend vers l'infini, la densité de la loi $t(n)$ est assimilée à la densité de la loi $\mathcal{N}(0; 1)$.

3. Loi de Fisher $F(n; m)$

Soient $n \in \mathbb{N}^*$ et $m \in \mathbb{N}^*$. La variable aléatoire X , à valeurs positives, suit la loi de **Fisher** à n et m degrés de liberté, notée $F(n; m)$, si X a pour densité :

$$\forall x \in \mathbb{R}, \quad f_X(x) = \frac{\left(\frac{n}{m}\right)^{\frac{n}{2}}}{B\left(\frac{n}{2}, \frac{m}{2}\right)} \times \frac{x^{\frac{n}{2}-1}}{\left(1 + \frac{n}{m}x\right)^{\frac{n+m}{2}}} \mathbb{I}_{x \geq 0},$$

où $B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} = \int_0^1 x^{a-1}(1-x)^{b-1} dx$ est la fonction Beta d'Euler.

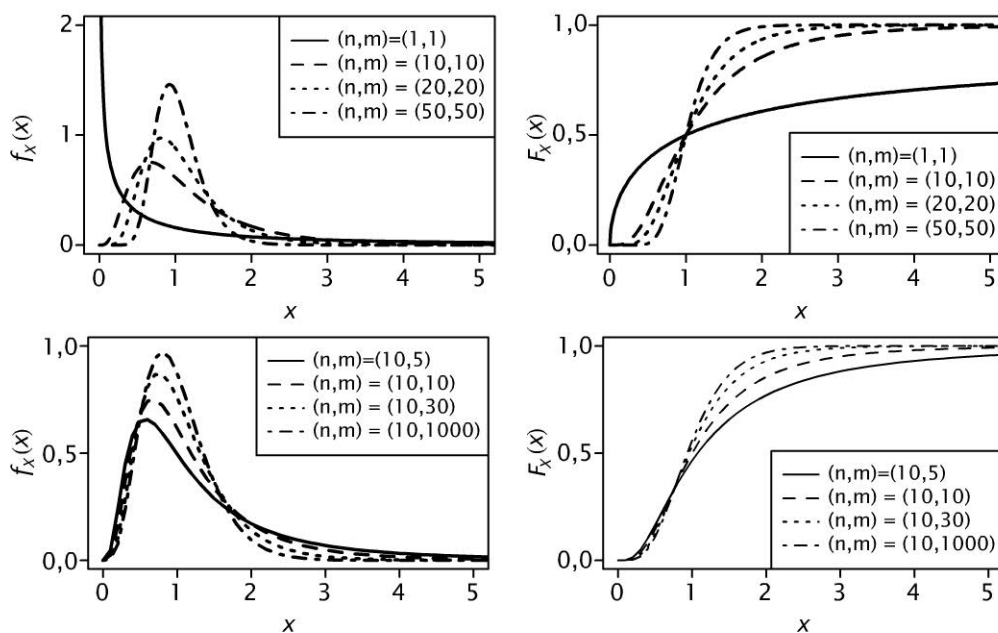


Figure 85.3 – Loi $F(n; m)$ pour différents choix de n et m : densité à gauche, fonction de répartition à droite.

Théorème 85.1

Si F suit la loi de Fisher de degrés de liberté n et m , alors $\frac{1}{F}$ suit la loi de Fisher de degrés de liberté m et n .

Dans le service des urgences d'un hôpital, l'équipe du matin commence à 7 h. L'instant du $k^{\text{ème}}$ appel (en minutes), noté T_k , suit une loi gamma de paramètres k et 1. Nous posons $T_0 = 0$ (qui correspond à l'arrivée de l'équipe du matin). Montrer qu'il est possible, à l'aide de la table de la loi du Khi-deux, de déterminer la probabilité qu'au bout d'un quart d'heure, l'équipe ait reçu n appels, pour $n \in \mathbb{N}$. Calculer cette probabilité pour $n = 10$. Nous noterons N_t le nombre d'appels reçus aux urgences, t minutes après l'arrivée de l'équipe du matin.

1. Posons $t = 15$. Nous nous intéressons à l'événement $\{N_{15} \geq n\}$.



Cherchons la relation entre N_t le nombre d'appels reçus pendant l'intervalle de temps $[0; t]$ et les variables T_k , $k \in \mathbb{N}^*$.

Nous avons

$$N_t = \sum_{k=1}^{+\infty} \mathbb{I}_{T_k \leq t}.$$

De plus les événements $\{N_t \geq n\}$ et $\{T_n \leq t\}$ sont les mêmes : ils signifient tous deux qu'au moins n appels ont été reçus pendant l'intervalle de temps $[0; t]$, ce qui signifie également que le $n^{\text{ème}}$ appel a été reçu avant l'instant t .

2. Compte tenu de la remarque précédente, nous avons donc

$$\mathbb{P}(N_t \geq n) = \mathbb{P}(T_n \leq t) = \mathbb{P}(2T_n \leq 2t) = F_{2T_n}(2t),$$

où F_{2T_n} est la fonction de répartition de la variable aléatoire $2T_n$.



Nous pouvons remarquer que $2T_n$ suit la loi du Khi-deux à $2n$ degrés de liberté. En effet,

$$F_{2T_n}(x) = \mathbb{P}(2T_n \leq x) = \mathbb{P}\left(T_n \leq \frac{x}{2}\right) = \int_0^{\frac{x}{2}} \frac{1}{\Gamma(n)} t^{n-1} \exp(-t) dt = \int_0^x \frac{1}{2^n \Gamma(n)} z^{n-1} \exp\left(-\frac{z}{2}\right) dz$$

par le changement de variable $z = 2t$.

Nous reconnaissons la fonction de répartition d'une loi du Khi-deux à $2n$ degrés de liberté.

La probabilité cherchée peut donc se lire dans la table du Khi-deux à $2n$ degrés de liberté.

3. Déterminons $\mathbb{P}(N_t \geq 10) = F_{2T_{10}}(2t)$. Nous prenons la table du Khi-deux à 20 degrés de liberté, et nous cherchons la probabilité qu'une variable aléatoire de loi $\chi^2(20)$ soit inférieure à $2t$ c'est-à-dire 30. D'après la table, cette probabilité est entre 0,90 et 0,95. Avec un logiciel de calcul (le langage **R** par exemple), nous obtenons une valeur approchée de la probabilité : celle-ci vaut 0,93.

Pour s'entraîner (solutions p. 547)

85.1 Soit X une variable aléatoire suivant une loi du Khi-deux à 9 degrés de liberté. Déterminer

1. x_1 tel que $\mathbb{P}(X \leq x_1) = 1 \%$,
2. x_2 tel que $\mathbb{P}(X \geq x_2) = 5 \%$,
3. $\mathbb{P}(X \geq 4,1682)$.

85.2 Soit T une variable aléatoire suivant une loi de Student à 15 degrés de liberté. Déterminer

1. t_1 tel que $\mathbb{P}(T \leq -t_1) + \mathbb{P}(T \geq t_1) = 5 \%$;
2. t_2 tel que $\mathbb{P}(T \geq t_2) = 10 \%$;

3. t_3 tel que $\mathbb{P}(T \leq t_3) = 5\%$;

4. $\mathbb{P}(T \geq 2,1314)$;

5. $\mathbb{P}(T \leq -2,1314)$.

85.3 Soit F une variable aléatoire suivant une loi de Fisher à 6 et 4 degrés de liberté. Déterminer

1. f_1 tel que $\mathbb{P}(F \leq f_1) = 5\%$;

2. f_2 tel que $\mathbb{P}(F \geq f_2) = 5\%$;

3. $\mathbb{P}(F \geq 9,1973)$.

85.4 Montrer que la loi du Khi-deux est un cas particulier de la loi gamma.

85.5 Soit X une variable aléatoire dont la densité est la densité de Cauchy, c'est-à-dire :

$\forall x \in \mathbb{R}, f_X(x) = \frac{1}{\pi} \times \frac{1}{1+x^2}$. Montrer que la loi de Cauchy est un cas particulier de la loi de Student.

85.6 Soit T une variable aléatoire de loi de Student à n degrés de liberté.

1. Quelle est la loi de $-T$?

2. Quelle est la loi de T^2 ?

85.7 

Soit F une variable aléatoire de loi de Fisher, de degrés de liberté n et m . Montrer que $1/F$ suit la loi de Fisher de degrés de liberté n et m .

Indication : Chercher la fonction de répartition de $X = 1/F$, puis la densité de X .

Couple aléatoire

Loi d'un couple discret

Les phénomènes aléatoires sont souvent décrits par plusieurs variables aléatoires réelles. Par exemple, dans l'étude d'une centrale nucléaire, nous notons X_1, X_2, X_3 les taux de radioactivité à 100 m (X_1), à 1 km de la centrale (X_2), et la température de l'eau de refroidissement (X_3). Dans cet exemple, (X_1, X_2, X_3) est appelé **vecteur aléatoire**. Dans cette Fiche nous étudions les vecteurs aléatoires de dimension deux : les **couples aléatoires**. Les définitions et propriétés données ici pour les couples aléatoires se généralisent pour les vecteurs aléatoires de dimension supérieure à deux.

Définition 86.1

Un **couple aléatoire** (X, Y) est une application de Ω dans $\mathbb{R}^2$, i.e. $(X, Y) :$

$$\begin{cases} \Omega \longrightarrow \mathbb{R}^2 \\ \omega \longmapsto (X(\omega); Y(\omega)) \end{cases}$$

Pour tout $\omega \in \Omega$, le couple de réels $(X(\omega), Y(\omega))$ est une **réalisation** du couple aléatoire (X, Y) . L'ensemble $X(\Omega) \times Y(\Omega)$ est l'ensemble des **valeurs prises par le couple** (X, Y) .



Il n'est pas nécessaire de spécifier Ω . Seul l'ensemble $X(\Omega) \times Y(\Omega)$ est indispensable.

Fiche 75

Exemple

Considérons une centrale nucléaire.

- Notons $X = 1$ s'il y a de l'eau dans le bassin de refroidissement, $X = 0$ sinon. Notons $Y = 1$ si la température de l'eau de refroidissement permet de refroidir suffisamment l'énergie thermique de la centrale, $Y = 0$ sinon. Nous avons $X(\Omega) \times Y(\Omega) = \{0; 1\} \times \{0; 1\}$. Le couple (X, Y) est **discret**.
- Notons X la quantité d'énergie thermique non transformée en énergie électrique, et Y la température de l'eau de refroidissement de la centrale. Nous avons $X(\Omega) \times Y(\Omega) = [0; +\infty[\times [0; +\infty[$. Le couple (X, Y) est **continu**.
- Notons X la quantité d'énergie thermique non transformée en énergie électrique, et $Y = 1$ si la température de l'eau de refroidissement de la centrale est suffisamment basse, $Y = 0$ sinon. Nous avons $X(\Omega) \times Y(\Omega) = [0; +\infty[\times \{0; 1\}$. Le couple (X, Y) est **mixte**.

Définissons maintenant la loi d'un couple aléatoire discret.

Définition 86.2

Soient X, Y deux variables aléatoires réelles. Si $X(\Omega)$ et $Y(\Omega)$ sont finis ou dénombrables, alors (X, Y) est un couple aléatoire **discret**. Dans ce cas, la **loi** de (X, Y) est définie par la famille de réels $(p_{X,Y}(x, y))_{(x,y) \in X(\Omega) \times Y(\Omega)}$ compris entre 0 et 1 satisfaisant

$\sum_{(x,y) \in X(\Omega) \times Y(\Omega)} p_{X,Y}(x,y) = 1$ et telle que pour tout événement $A \subset X(\Omega) \times Y(\Omega)$, la probabilité que (X, Y) appartienne à A est

$$\mathbb{P}((X, Y) \in A) = \sum_{(x,y) \in A} p_{X,Y}(x,y).$$

Les **lois marginales de X et Y** sont définies par :

$$\forall x \in X(\Omega), \quad p_X(x) = \mathbb{P}(X = x) = \sum_{y \in Y(\Omega)} p_{X,Y}(x,y),$$

$$\forall y \in Y(\Omega), \quad p_Y(y) = \mathbb{P}(Y = y) = \sum_{x \in X(\Omega)} p_{X,Y}(x,y).$$

Exemple

Dans un jeu de 32 cartes, tirons deux cartes au hasard. Soient X et Y les nombres de rois et de dames tirés. La loi de probabilité de (X, Y) et les lois marginales sont données dans le tableau ci-dessous. La probabilité de tirer au moins un roi est $\mathbb{P}(X = 1) + \mathbb{P}(X = 2) = p_{X,Y}(1; 0) + p_{X,Y}(1; 1) + p_{X,Y}(2; 0)$.

$p_{X,Y}(x,y)$	$x = 0$	$x = 1$	$x = 2$	$p_Y(y) = \mathbb{P}(Y = y)$
$y = 0$	C_{24}^2/C_{32}^2	$C_4^1 C_{24}^1/C_{32}^2$	C_4^2/C_{32}^2	C_{28}^2/C_{32}^2
$y = 1$	$C_4^1 C_{24}^1/C_{32}^2$	$C_4^1 C_4^1/C_{32}^2$	0	$C_4^1 C_{28}^1/C_{32}^2$
$y = 2$	C_4^2/C_{32}^2	0	0	C_4^2/C_{32}^2
$p_X(x) = \mathbb{P}(X = x)$	C_{28}^2/C_{32}^2	$C_4^1 C_{28}^1/C_{32}^2$	C_4^2/C_{32}^2	1

Les coefficients C_n^{2k} ont été définis à la Fiche 71.

 Fiche 71



- Dans l'exemple précédent, X et Y ont la même loi, mais ne sont pas égaux (l'un est le nombre de rois, l'autre le nombre de dames).
- Un couple (x, y) pour lequel $p_{X,Y}(x, y)$ est proche de 1 a une grande probabilité d'apparaître, tandis que si $p_{X,Y}(x, y)$ est proche de 0, le couple (x, y) a une faible probabilité d'apparaître.

Application La loi multinomiale

Une expérience est réalisée m fois en laboratoire, et a trois issues possibles : succès, échec, ou résultat non interprétable (avec probabilités respectives p_s, p_e, p_n , tels que $p_s + p_e + p_n = 1$). Notons X et Y les nombres de succès et d'échecs. Montrer que la loi de (X, Y) est donnée par

$$p_{X,Y}(x,y) = \frac{m!}{x!y!(m-x-y)!} p_s^x p_e^y p_n^{m-x-y} \text{ pour tous } x, y \in \{0; \dots; m\} \text{ tels que } x + y \leq m.$$

Cette loi porte le nom de **loi multinomiale** de paramètres $(m; p_s; p_e)$.

Montrer que les lois de X et Y sont des lois binomiales de paramètres respectifs $(m; p_s)$ et $(m; p_e)$.

Calculons la probabilité $\mathbb{P}(X = x, Y = y)$, avec x, y des entiers.

1. Bien sûr, si $x + y > m$, la probabilité cherchée est nulle.

2. Considérons x, y deux entiers tels que $x + y \leq m$. Alors, $\mathbb{P}(X = x, Y = y)$ est la probabilité d'obtenir x succès et y échecs parmi les m expériences. C'est aussi la probabilité d'obtenir $x + y$ résultats interprétables, dont x succès. Notons $R_i = X + Y$ le nombre de résultats interprétables. Alors

$$\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x, R_i = x + y).$$

Ensuite, remarquons qu'il est assez facile de calculer $\mathbb{P}(X = x | R_i = x + y)$. En effet, il s'agit de calculer la probabilité que, parmi $x + y$ expériences interprétables, il y en ait x qui soient un succès.



Notons S l'événement « le résultat de l'expérience est un succès », E l'événement « le résultat de l'expérience est un échec », et I l'événement « le résultat de l'expérience est interprétable ». Nous avons $I = S \cup E$, $\mathbb{P}(S) = p_s$, $\mathbb{P}(E) = p_e$ et $\mathbb{P}(I) = \mathbb{P}(S \cup E) = p_s + p_e$ car les événements S et E sont disjoints.

a. Calculons $\mathbb{P}(X = x | R_i = x + y)$. Nous avons $x + y$ expériences interprétables. Toutes ces expériences interprétables sont réalisées indépendamment les unes des autres, et pour chacune de ces expériences interprétables, notons p la probabilité qu'elle soit un succès. Nous reconnaissons ici une loi binomiale de paramètres $x + y$ et p . Calculons p .



Attention, p n'est pas p_s ici :

- p_s est la probabilité qu'une expérience soit un succès : $p_s = \mathbb{P}(S)$,
- p est la probabilité qu'une expérience interprétable soit un succès : $p = \mathbb{P}(S | I)$.

Donc $p = \mathbb{P}(S | I) = \frac{\mathbb{P}(S \cap I)}{\mathbb{P}(I)} = \frac{\mathbb{P}(S)}{\mathbb{P}(I)}$ car nous avons $S \cap I = S$ puisque $S \subset I$. Finalement,

$$p = \frac{p_s}{p_s + p_e}$$

et

$$\mathbb{P}(X = x | R_i = x + y) = \binom{x+y}{x} p^x (1-p)^y = \frac{(x+y)!}{x! y!} \times \frac{p_s^x}{(p_s + p_e)^x} \times \frac{p_e^y}{(p_s + p_e)^y}.$$

b. Calculons maintenant $\mathbb{P}(R_i = x + y)$. Nous remarquons ici que R_i suit une loi binomiale de paramètres $(m; p_s + p_e)$. En effet, m expériences sont réalisées, et chacun des résultats est interprétable avec une probabilité $p_s + p_e = 1 - p_n$. Nous avons donc

$$\mathbb{P}(R_i = x + y) = \binom{m}{x+y} (p_s + p_e)^{x+y} (1 - p_s - p_e)^{m-x-y}.$$

Pour conclure, il suffit d'utiliser la définition de la probabilité conditionnelle :

$$\begin{aligned}\mathbb{P}(X = x, R_i = x + y) &= \mathbb{P}(X = x | R_i = x + y) \mathbb{P}(R_i = x + y) \\ &= \frac{(x + y)!}{x! y!} \times \frac{p_s^x}{(p_s + p_e)^x} \times \frac{p_e^y}{(p_s + p_e)^y} \times \\ &\quad \times \binom{m}{x + y} \times (p_s + p_e)^{x+y} \times (1 - p_s - p_e)^{m-x-y} \\ &= \frac{m!}{x! y! (m - x - y)!} \times p_s^x \times p_e^y \times (1 - p_s - p_e)^{m-x-y}.\end{aligned}$$

3. Calculons maintenant la loi de X . Par définition de la loi marginale, nous avons :

$$\begin{aligned}\mathbb{P}(X = x) &= \sum_{y \in y(\Omega)} \mathbb{P}(X = x, Y = y) \\ &= \sum_{y=0}^{m-x} \frac{m!}{x! y! (m - x - y)!} \times p_s^x \times p_e^y \times p_n^{m-x-y} = \binom{m}{x} \times p_s^x \times (p_e + p_r)^{m-x} \\ \text{car } \sum_{y=0}^{m-x} \frac{(m-x)!}{y! (m-x-y)!} \times p_e^y \times p_n^{m-x-y} &= (p_e + p_n)^{m-x} \text{ par la formule du binôme de Newton, et} \\ \binom{m}{x} &= \frac{m!}{x! (m-x)!}.\end{aligned}$$

Nous reconnaissons que X suit la loi binomiale de paramètres $(m; p_s)$. De même, Y suit une loi binomiale de paramètres $(m; p_e)$.

Pour s'entraîner (solutions p. 548)

86.1 Un campus universitaire est constitué de trois écoles, chacune se partage le tiers des étudiants du campus. Nous supposons qu'il n'y a pas de possibilité d'être étudiant(e) dans plus d'une école à la fois. Dans l'école 1, le taux de filles est de 80 %. Dans les écoles 2 et 3, il est respectivement de 20 % et 50 %. Soit un(e) étudiant(e) rencontré(e) au hasard dans le campus. Notons $X = 1$ si l'étudiant(e) est une fille, 0 sinon ; et $Y = 1, 2$ ou 3 suivant que l'étudiant(e) est dans l'école 1, 2 ou 3.

1. Quelle est la loi de (X, Y) ?
2. Quelles sont les lois marginales ?
3. Pour une étudiante du campus, quelle est la probabilité qu'elle soit de l'école 2 ?

86.2 Un agriculteur a un troupeau de N bovins : b bœufs, ℓ vaches laitières, v veaux,

avec $b + \ell + v = N$. Deux bovins se sont échappés du pré. Notons $X = 1$ si le premier bovin échappé est un bœuf, 0 sinon. De même, notons $Y = 1$ si le second bovin est un bœuf, 0 sinon.

1. Quelle est la loi de (X, Y) ?
2. Quelles sont les lois marginales ?
3. Quelle est la loi de $Z = XY$? Reconnaitre une loi discrète connue.

86.3 Soit (X, Y) un couple aléatoire à valeurs dans $\mathbb{N}^* \times \mathbb{N}^*$ et tel que $\mathbb{P}(X = x, Y = y) = a \left(\frac{1}{2}\right)^{x+y}$, pour tout $(x, y) \in \mathbb{N}^* \times \mathbb{N}^*$.

1. Que vaut la constante a ?
2. Quelles sont les lois marginales ? Reconnaitre une loi discrète connue.

Loi d'un couple aléatoire continu

Dans la Fiche 86, nous avons défini la loi d'un couple aléatoire discret. Nous définissons maintenant la loi d'un couple aléatoire continu.

Définition 87.1

Soient X, Y deux variables aléatoires réelles. Si $X(\Omega)$ et $Y(\Omega)$ sont des intervalles de $\mathbb{R}$, nous disons que (X, Y) est un couple aléatoire **continu**. Dans ce cas, la loi de (X, Y) est définie par la **densité de probabilité** sur $\mathbb{R}^2$, c'est-à-dire une fonction $f_{X,Y}$ définie sur $\mathbb{R}^2$ et à valeurs dans $\mathbb{R}$ satisfaisant

$$1. \forall (x, y) \in \mathbb{R}^2, \quad f_{X,Y}(x, y) \geq 0,$$

$$2. \iint_{\mathbb{R}^2} f_{X,Y}(x, y) dx dy = 1,$$

et telle que pour tout événement $A \subset \mathbb{R}^2$, la probabilité que (X, Y) appartienne à A est égale à

$$\mathbb{P}((X, Y) \in A) = \iint_A f_{X,Y}(x, y) dx dy.$$

La loi de X est appelée **loi de probabilité de densité** $f_{X,Y}$. Les **densités marginales de X et Y** sont définies par :

$$\forall (x, y) \in \mathbb{R}^2, \quad f_X(x) = \int_{\mathbb{R}} f_{X,Y}(x, y) dy, \quad f_Y(y) = \int_{\mathbb{R}} f_{X,Y}(x, y) dx.$$

Exemple

Soit (X, Y) un couple aléatoire continu, à valeurs dans $\mathbb{R}^2$, et de densité $f_{X,Y}(x, y) = \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}}$ représentée sur le graphe de la Figure 87.1. Les zones où la densité est importante ont une grande probabilité d'apparaître, tandis que les zones où la densité est faible ont une faible probabilité d'apparaître. Dans le cas de la Figure 87.1, $\mathbb{P}((X, Y) \in [-0, 5; 0, 5] \times [-0, 5; 0, 5])$ est grande, tandis que $\mathbb{P}((X, Y) \in [1; 2] \times [1; 2])$ est faible.

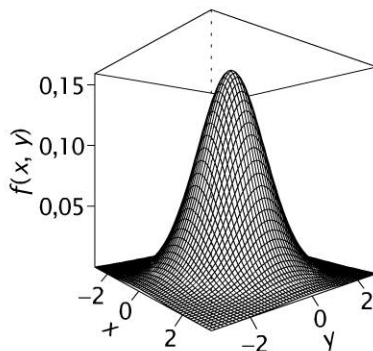


Figure 87.1 – Loi de couples aléatoires : densité $f(x, y)$ pour le cas continu.

Nous jetons un point au hasard dans le disque unité, et nous notons (X, Y) les coordonnées de ce point. Quelle est la densité de (X, Y) ? Quelle est la probabilité que le point arrive à une distance a de l'origine, avec $a > 0$? Montrer que X et Y ont la même loi, mais ne sont pas égaux.

1. Remarquons que le point ne peut pas tomber hors du disque, donc pour toute zone B du plan, située en dehors du disque, $\mathbb{P}((X, Y) \in B) = 0$. Or, $\mathbb{P}((X, Y) \in B) = \iint_B f_{X,Y}(x, y) dx dy$. Nous en concluons que la densité $f_{X,Y}$ est nulle en dehors du disque. Donc $f_{X,Y}$ sera de la forme

$$f_{X,Y}(x, y) = \dots \mathbb{1}_{(x,y) \in D}$$

où D est le disque unité. Nous rappelons que $\mathbb{1}_{(x,y) \in D}$ est égal à 1 si $(x, y) \in D$ et 0 si $(x, y) \notin D$.

2.



Ensuite, comme le point est lancé au hasard, chaque zone du disque unité a la même probabilité d'être atteinte. Nous allons donc certainement trouver une densité qui sera constante sur tout le disque, c'est-à-dire de la forme

$$f_{X,Y}(x, y) = C \mathbb{1}_{(x,y) \in D}.$$

Montrons-le correctement.

Considérons une zone B du plan, incluse dans le disque. Alors, comme le point est lancé au hasard, d'après la remarque suivant la Définition 73.5, nous avons $\mathbb{P}((X, Y) \in B) = \frac{\text{Aire de } B}{\pi}$ où π est l'aire du disque unité. Or, nous savons que

$$\mathbb{P}((X, Y) \in B) = \iint_B f_{X,Y}(x, y) dx dy.$$

La densité $f_{X,Y}$ cherchée doit donc vérifier

$$\iint_B f_{X,Y}(x, y) dx dy = \frac{\text{Aire de } B}{\pi}. \quad (87.1)$$

La fonction $f_{X,Y}$ qui convient ici est donc une fonction constante sur le disque :

$$f_{X,Y}(x, y) = C \mathbb{1}_{(x,y) \in D},$$

et il nous reste à trouver la constante C . Pour cela, remarquons que

$$\iint_B f_{X,Y}(x, y) dx dy = \iint_B C \mathbb{1}_{(x,y) \in D} dx dy = C \times \text{Aire de } B$$

ce qui conduit à $C = \frac{1}{\pi}$, grâce à l'Équation (87.1).

Signalons que nous pouvons calculer C en utilisant la propriété 2 de la Définition 87.1 d'une densité. En effet,

$$1 = \iint_{\mathbb{R}^2} f_{X,Y}(x, y) dx dy = \iint_{\mathbb{R}^2} C \mathbb{1}_{(x,y) \in D} dx dy = C \iint_D dx dy = \frac{C}{\pi}$$

d'où encore $C = \frac{1}{\pi}$.

Fiche 73

3. Calculons maintenant la probabilité que le point tombe à une distance a de l'origine. Notons D_a le disque centré en l'origine, et de rayon a . Alors la probabilité que le point tombe à une distance a de l'origine est $\mathbb{P}((X, Y) \in D_a)$.

- Si $a > 1$, cette probabilité est 1, car le point est jeté dans le disque de rayon 1.
- Si $0 < a \leq 1$, nous avons $\mathbb{P}((X, Y) \in D_a) = \iint_{D_a} \frac{1}{\pi} dx dy = \frac{1}{\pi} \times \text{Aire de } D_a = \frac{1}{\pi} \times \pi a^2 = a^2$.

Pour tout $a > 0$, nous pouvons conclure que la probabilité que le point tombe à une distance a de l'origine est :

$$\mathbb{P}((X, Y) \in D_a) = a^2 \mathbb{I}_{0 < a \leq 1} + \mathbb{I}_{a > 1}.$$

4. Enfin, calculons les lois marginales. Commençons par la loi de X :

$$f_X(x) = \int_{-\infty}^{+\infty} f_{X,Y}(x, y) dy = \begin{cases} 0 & \text{si } x < -1 \\ 0 & \text{si } x > 1 \\ \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \frac{1}{\pi} dy & \text{si } -1 \leq x \leq 1 \end{cases}$$

c'est-à-dire

$$f_X(x) = \frac{2}{\pi} \sqrt{1-x^2} \mathbb{I}_{x \in [-1; 1]}.$$

Par symétrie de x et y dans la loi jointe $f_{X,Y}(x, y)$, nous avons aussi

$$f_Y(y) = \frac{2}{\pi} \sqrt{1-y^2} \mathbb{I}_{y \in [-1; 1]}.$$

5. X et Y ont la même densité, c'est-à-dire qu'elles ont la même loi. Mais attention, cela ne signifie pas qu'elles soient égales. En effet, nous avons

$$\mathbb{P}((X, Y) \in B_1) + \mathbb{P}((X, Y) \in B_2) + \mathbb{P}((X, Y) \in B_3) = 1 \quad (87.2)$$

$$B_1 = \{(x, y) \in D \text{ tels que } x > y\}$$

où les ensembles $B_2 = \{(x, y) \in D \text{ tels que } x < y\}$ sont représentés sur la Figure 87.2.

$$B_3 = \{(x, y) \in D \text{ tels que } x = y\}$$

Le lancer du point étant uniforme, nous avons $\mathbb{P}((X, Y) \in B_1) = \mathbb{P}((X, Y) \in B_2) = \frac{1}{2}$. Nous déduisons de (87.2) que $\mathbb{P}((X, Y) \in B_3) = 0$, c'est-à-dire $\mathbb{P}(X = Y) = 0$, c'est-à-dire X et Y ne sont pas égaux.

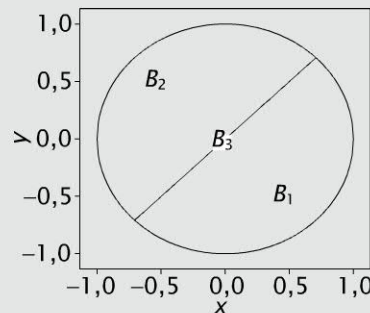


Figure 87.2 – Ensembles B_1 , B_2 et B_3

Pour s'entraîner (solutions p. 548)

87.1 Soit (X, Y) un couple aléatoire de densité f définie sur $\mathbb{R}^2$. Exprimer en fonction de f , la probabilité $\mathbb{P}(a < X \leq b, c < Y \leq d)$.

87.2 Calculer les lois marginales de la densité $f_{X,Y}$:

$$\forall (x, y) \in \mathbb{R}^2, \quad f_{X,Y}(x, y) = \frac{1}{2\pi} e^{-\frac{x^2 + y^2}{2}}.$$

87.3 Un point est jeté au hasard dans le carré de côté 1, quelle est la densité du couple (X, Y) ? Quelles sont les lois marginales ?

87.4 

Soit la fonction définie sur $\mathbb{R}^2$ par

$$f(x, y) = C(x + y - 2xy) \mathbb{1}_{(x,y) \in [0;1] \times [0;1]}.$$

1. Que vaut la constante C pour que f soit une densité d'un couple aléatoire (X, Y) ?

2. Quelle est la probabilité $\mathbb{P}(|X - Y| \leq \frac{1}{2})$?

87.5 Soit $f(x, y) = C e^{-2x} e^{-y} \mathbb{1}_{(x,y) \in \mathbb{R}^+ \times \mathbb{R}^+}$.

1. Que vaut la constante C pour que f soit une densité d'un couple aléatoire (X, Y) ?

2. Calculer les lois marginales. Reconnaître des lois connues.

3. Calculer les probabilités suivantes :

(a) $\mathbb{P}(X < 1, Y > 1)$;

(b) $\mathbb{P}(X > Y)$;

(c) $\mathbb{P}(X < 1)$.

Fonction de répartition d'un couple aléatoire

Nous avons vu que, pour les variables aléatoires réelles, la loi est caractérisée par la fonction de répartition. C'est encore le cas des couples aléatoires : la **loi d'un couple aléatoire** est déterminée par la **fonction de répartition conjointe** définie ci-dessous.

Définition 88.1

Soit (X, Y) un couple aléatoire. Sa **fonction de répartition conjointe** est la fonction $F_{X,Y}$ définie par :

$$F_{X,Y} : \begin{cases} \mathbb{R}^2 & \longrightarrow [0; 1] \\ (x, y) & \longmapsto F_{X,Y}(x, y) = \mathbb{P}(X \leq x, Y \leq y) \end{cases} .$$

Théorème 88.1

Soient (X, Y) un couple de variables aléatoires réelles et $F_{X,Y}$ sa fonction de répartition. La fonction $(x, y) \mapsto F_{X,Y}(x, y)$ est croissante par rapport à x et y , et

$$F_{X,Y}(-\infty, y) = F_{X,Y}(x, -\infty) = F_{X,Y}(-\infty, -\infty) = 0, \quad F_{X,Y}(+\infty, +\infty) = 1.$$

Définissons maintenant les **fonctions de répartition marginales**.

Définition 88.2

Soit (X, Y) un couple aléatoire. Les fonctions de répartition F_X et F_Y sont appelées **fonctions de répartition marginales** de X et Y . Elles caractérisent les **lois** de X et Y . La fonction de répartition de X se déduit de la fonction de répartition conjointe de (X, Y) :

$$\forall x \in \mathbb{R}, \quad F_X(x) = \mathbb{P}(X \leq x) = \mathbb{P}(X \leq x, Y < +\infty) = \lim_{y \rightarrow +\infty} F_{X,Y}(x, y) = F_{X,Y}(x, +\infty).$$

De même, la fonction de répartition de Y est :

$$\forall y \in \mathbb{R}, \quad F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(X < +\infty, Y \leq y) = \lim_{x \rightarrow +\infty} F_{X,Y}(x, y) = F_{X,Y}(+\infty, y).$$



Quand la loi d'un couple aléatoire (X, Y) est connue, nous pouvons en déduire ses lois marginales, mais la réciproque est fautive. En effet, les lois marginales n'expliquent pas quel est le type de relation liant X et Y , tandis que la loi jointe l'explique.

Exemple

Soit X la valeur d'une action boursière le jour j et Y la valeur de la même action le jour $j + 1$. Supposons que l'économie est stable sur ces deux jours. Nous pouvons faire l'hypothèse que

X et Y ont même loi, autrement dit que F_X et F_Y sont identiques. Par contre, si nous voulons calculer la probabilité que l'action vaille moins d'un euro le jour j et qu'elle vaille strictement plus de 100 euros le jour $j + 1$, nous avons besoin de connaître $F_{X,Y}$. En effet, cette probabilité est $\mathbb{P}(X \leq 1, Y > 100) = \mathbb{P}(X \leq 1) - \mathbb{P}(X \leq 1, Y \leq 100)$, et ceci est égal à $F_X(1) - F_{X,Y}(1; 100)$.

Nous présentons maintenant la forme de la fonction de répartition dans les cas de couples aléatoires discret et continu.

Théorème 88.2

Soit (X, Y) un couple aléatoire **discret**, dont la loi est définie par la famille de réels $(p_{X,Y}(x, y))_{(x,y) \in X(\Omega) \times Y(\Omega)}$. La **fonction de répartition conjointe** $F_{X,Y}$ du couple (X, Y) est définie par :

$$\forall (x, y) \in X(\Omega) \times Y(\Omega), \quad F_{X,Y}(x, y) = \sum_{u \leq x, v \leq y} p_{X,Y}(u, v).$$

Les lois marginales de X et Y sont définies par :

$$\forall x \in X(\Omega), \quad p_X(x) = \mathbb{P}(X = x) = \sum_{y \in Y(\Omega)} p_{X,Y}(x, y),$$

$$\forall y \in Y(\Omega), \quad p_Y(y) = \mathbb{P}(Y = y) = \sum_{x \in X(\Omega)} p_{X,Y}(x, y).$$

Et les fonctions de répartition marginales sont données par

$$\forall (x, y) \in X(\Omega) \times Y(\Omega), \quad F_X(x) = \sum_{u \leq x} p_X(u), \quad F_Y(y) = \sum_{v \leq y} p_Y(v).$$

Théorème 88.3

Soit (X, Y) un couple aléatoire **continu**, dont la loi est déterminée par la densité de probabilité $(x, y) \mapsto f_{X,Y}(x, y)$ définie sur $\mathbb{R}^2$. La **fonction de répartition** de (X, Y) est reliée à $f_{X,Y}$ par la relation

$$\forall (x, y) \in \mathbb{R}^2, \quad F_{X,Y}(x, y) = \int_{-\infty}^x \int_{-\infty}^y f_{X,Y}(u, v) du dv.$$

Si $F_{X,Y}$ est de classe $\mathcal{C}^2$ en $(x, y) \in \mathbb{R}^2$ alors

$$f_{X,Y}(x, y) = \frac{\partial^2 F_{X,Y}(x, y)}{\partial x \partial y}.$$

Les lois marginales de X et Y sont définies par :

$$\forall x \in \mathbb{R}, \quad f_X(x) = \int_{-\infty}^{+\infty} f_{X,Y}(x, y) dy,$$

$$\forall y \in \mathbb{R}, \quad f_Y(y) = \int_{-\infty}^{+\infty} f_{X,Y}(x, y) dx.$$

De même, les fonctions de répartition marginales sont données par

$$\forall (x, y) \in \mathbb{R}^2, \quad F_X(x) = \int_{-\infty}^x f_X(u) du, \quad F_Y(y) = \int_{-\infty}^y f_Y(v) dv.$$

Soient (X, Y) les coordonnées d'un point jeté au hasard dans le carré de côté 2, et de sommets $(0; 0)$, $(2; 0)$, $(2; 2)$ et $(0; 2)$.

1. Quelle est la fonction de répartition de (X, Y) ?
2. Quelles sont les fonctions de répartition marginales ?
3. Quelle est la probabilité que le point tombe dans le carré de sommets $(0; 0)$, $(1; 0)$, $(1; 1)$ et $(0; 1)$?

1. Le point est jeté dans le carré de côté 2, il ne peut donc pas tomber à gauche de la droite $x = 0$, donc $F_{X,Y}(x, y) = 0$ si $x < 0$. Par le même raisonnement, $F_{X,Y}(x, y) = 0$ si $y < 0$, donc

$$F_{X,Y}(x, y) = 0 \quad \text{si } x < 0 \text{ ou si } y < 0.$$

De même, le point tombe nécessaire à gauche de la droite $x = 2$ et en dessous de la droite $y = 2$; donc

$$F_{X,Y}(x, y) = 1 \quad \text{si le couple } (x, y) \text{ vérifie } x \geq 2 \text{ et } y \geq 2.$$

Distinguons les trois autres cas.

- a. Si $y \geq 2$ et $0 \leq x \leq 2$,

$$F_{X,Y}(x, y) = \mathbb{P}(X \leq x) = \frac{A_x}{4} = \frac{2x}{4} = \frac{x}{2}$$

où 4 est l'aire du carré de côté 2 et A_x est l'aire de l'ensemble D_x représenté dans la Figure 88.1.

- b. Si $x \geq 2$ et $0 \leq y \leq 2$,

$$F_{X,Y}(x, y) = \mathbb{P}(Y \leq y) = \frac{A_y}{4} = \frac{2y}{4} = \frac{y}{2}$$

où A_y est l'aire de l'ensemble D_y représenté dans la Figure 88.1.

- c. Enfin, si $0 \leq x \leq 2$ et $0 \leq y \leq 2$, nous avons :

$$F_{X,Y}(x, y) = \mathbb{P}(X \leq x, Y \leq y) = \frac{A_{x,y}}{4} = \frac{xy}{4}$$

où $A_{x,y}$ désigne l'aire de l'ensemble $D_{x,y}$ représenté sur la Figure 88.1.

2. La fonction de répartition de X est donnée par

$$F_X(x) = F_{X,Y}(x, +\infty) = \begin{cases} 0 & \text{si } x \leq 0 \\ \frac{x}{2} & \text{si } x \in [0; 2] \\ 1 & \text{si } x \geq 2 \end{cases}.$$

De même, la fonction de répartition de Y est donnée par

$$F_Y(y) = F_{X,Y}(+\infty, y) = \begin{cases} 0 & \text{si } y \leq 0 \\ \frac{y}{2} & \text{si } y \in [0; 2] \\ 1 & \text{si } y \geq 2 \end{cases}.$$

X et Y ont donc la même loi (mais ne sont pas égaux ! D'ailleurs, nous avons $\mathbb{P}(X = Y) = 0$).

3. En considérant les mêmes notations qu'en 1., la probabilité que le point tombe dans le carré de côté 1 est $\mathbb{P}((X, Y) \in D_{1,1}) = \frac{A_{1,1}}{4} = \frac{1}{4}$.

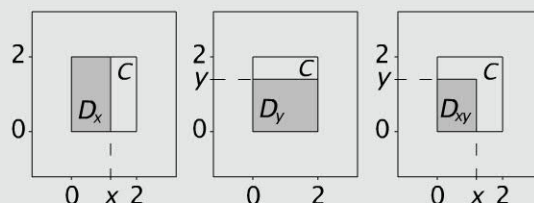


Figure 88.1

Pour s'entraîner (solutions p. 548)

88.1 Soit (X, Y) un couple aléatoire de fonction de répartition $F_{X,Y}$. Soient $a < b$ et $c < d$ des réels. Exprimer $\mathbb{P}(a < X \leq b, c < Y \leq d)$ en fonction de $F_{X,Y}$.

88.2 Notons X et Y les consommations électriques mensuelles (en kW) de deux appartements. Soient $U = \max(X, Y)$ et $V = \min(X, Y)$ les consommations maximale et minimale des deux appartements.

1. Exprimer les fonctions de répartition de U et V en fonction de $F_{X,Y}$ la fonction de répartition de (X, Y) .
2. Calculer la fonction de répartition de (U, V) . Vérifier que vous retrouvez bien les fonctions de répartition marginales trouvées à la question précédente.

3. Quelle est la probabilité que l'un des appartements consomme plus de 300 kW et l'autre moins de 100 kW ?

88.3

www

Soient (X, Y) les coordonnées d'un point jeté au hasard dans le disque unité.

1. Quelle est la fonction de répartition de (X, Y) ?
2. Quelles sont les fonctions de répartition marginales ?
3. Quelle est la probabilité que le point tombe dans le carré de côtés $(0; 0), \left(\frac{1}{2}; 0\right), \left(\frac{1}{2}; \frac{1}{2}\right), \left(0; \frac{1}{2}\right)$?

Moments de plusieurs variables aléatoires réelles

Nous avons déjà évoqué la nécessité de travailler avec plusieurs variables aléatoires réelles. Nous expliquons maintenant comment calculer une moyenne d'une fonction de plusieurs variables aléatoires réelles, par exemple la moyenne de $U = RI$ où R et I sont deux variables aléatoires.

 Fiche 86

1. Définition générale

Dans le cas d'une seule variable aléatoire réelle, nous avons défini l'espérance $\mathbb{E}(g(X))$ d'une fonction quelconque $g : \mathbb{R} \rightarrow \mathbb{R}$ de la variable aléatoire réelle X . À présent, nous définissons l'espérance $\mathbb{E}(g(X, Y))$ d'une fonction quelconque $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ du couple aléatoire (X, Y) .

 Fiche 77

Définition 89.1

Soient X et Y deux variables aléatoires réelles. L'**espérance** de $g(X, Y)$, notée $\mathbb{E}(g(X, Y))$ est définie par :

$$\mathbb{E}(g(X, Y)) = \sum_{x \in X(\Omega)} \sum_{y \in Y(\Omega)} g(x, y) \mathbb{P}(X = x, Y = y)$$

si le couple (X, Y) est discret, (89.1)

$$\mathbb{E}(g(X, Y)) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x, y) f_{X,Y}(x) dx dy$$

si le couple (X, Y) est à densité, (89.2)

où, si (X, Y) est un couple aléatoire à densité, $f_{X,Y}$ désigne sa densité jointe.



La définition de l'espérance suppose que les quantités (89.1) et (89.2) existent.

 Fiches 65 et 66

Exemple

D'après la définition précédente, nous avons :

$$\mathbb{E}(XY) = \sum_{x \in X(\Omega)} \sum_{y \in Y(\Omega)} xy \mathbb{P}(X = x, Y = y) \quad \text{si } (X, Y) \text{ est un couple aléatoire discret,}$$

$$\mathbb{E}(XY) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} xy f_{X,Y}(x) dx dy \quad \text{si } (X, Y) \text{ est un couple aléatoire à densité.}$$

2. Covariance, coefficient de corrélation et variance

Soient X et Y deux variables aléatoires réelles. Nous avons vu aux Fiches 77 et 78, que $\mathbb{E}(X)$, $\mathbb{E}(Y)$, $\text{Var}(X)$ et $\text{Var}(Y)$ donnent des informations sur X et Y séparément, mais aucune information sur la nature de la relation existant entre X et Y .

Une information sur cette relation est donnée par le **coefficient de corrélation**. Définissons tout d'abord la **covariance**.

Définition 89.2

La **covariance** de deux variables aléatoires réelles X et Y est définie par :

$$\mathbb{C}\text{ov}(X, Y) = \mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))),$$

c'est-à-dire

$$\mathbb{C}\text{ov}(X, Y) = \sum_{x \in X(\Omega)} \sum_{y \in Y(\Omega)} (x - \mathbb{E}(X))(y - \mathbb{E}(Y)) \mathbb{P}(X = x, Y = y)$$

si le couple (X, Y) est discret,

$$\mathbb{C}\text{ov}(X, Y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - \mathbb{E}(X))(y - \mathbb{E}(Y)) f_{X,Y}(x, y) dx dy$$

si le couple (X, Y) est à densité,

où, si (X, Y) est un couple aléatoire à densité, $f_{X,Y}$ désigne sa densité.

Nous pouvons également définir la **matrice de variance-covariance** du couple aléatoire (X, Y) par la matrice symétrique suivante :

$$C = \begin{pmatrix} \text{Var}(X) & \mathbb{C}\text{ov}(X, Y) \\ \mathbb{C}\text{ov}(X, Y) & \text{Var}(Y) \end{pmatrix}.$$

Théorème 89.1

La covariance a les **propriétés** suivantes :

- $\mathbb{C}\text{ov}(X, Y) = \mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$

c'est-à-dire

$$\mathbb{C}\text{ov}(X, Y) = \left(\sum_{x \in X(\Omega)} \sum_{y \in Y(\Omega)} xy \mathbb{P}(X = x, Y = y) \right) - \mathbb{E}(X)\mathbb{E}(Y) \quad \text{si } (X, Y) \text{ est discret,}$$

$$\mathbb{C}\text{ov}(X, Y) = \left(\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} xy f_{X,Y}(x, y) dx dy \right) - \mathbb{E}(X)\mathbb{E}(Y) \quad \text{si } (X, Y) \text{ est à densité.}$$

- $\mathbb{C}\text{ov}(X, Y) = \mathbb{C}\text{ov}(Y, X)$;
- $\mathbb{C}\text{ov}(X, X) = \text{Var}(X)$;
- $\mathbb{C}\text{ov}(aX + b, cY + d) = ac \mathbb{C}\text{ov}(X, Y)$ où a, b, c, d sont des réels ;
- $\mathbb{C}\text{ov}\left(\sum_{i=1}^n X_i, \sum_{j=1}^m Y_j\right) = \sum_{i=1}^n \sum_{j=1}^m \mathbb{C}\text{ov}(X_i, Y_j).$



La covariance entre X et Y est l'énergie croisée des deux variables. C'est un nombre réel (positif ou négatif). C'est une grandeur dont l'unité de mesure est le produit des unités de X et de Y . Nous préférons travailler avec une grandeur adimensionnelle : le **coefficient de corrélation**.

Définition 89.3

Le **coefficient de corrélation** de deux variables aléatoires réelles X et Y de variances non nulles est défini par

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)} \sqrt{\text{Var}(Y)}}.$$

Nous pouvons également définir la **matrice de corrélation** du couple aléatoire (X, Y) par la matrice symétrique suivante :

$$\rho = \begin{pmatrix} 1 & \rho(X, Y) \\ \rho(X, Y) & 1 \end{pmatrix}.$$

Théorème 89.2

Le coefficient de corrélation a les **propriétés** suivantes :

- $\rho(X, Y) = \rho(Y, X)$;
- $-1 \leq \rho(X, Y) \leq +1$;
- $\rho(X, X) = 1$.



- Une valeur de $\rho(X, Y)$ proche de $+1$ ou -1 indique une relation de linéarité très forte entre les variables X et Y .
- Une valeur de $\rho(X, Y)$ proche de 0 indique une absence de toute relation linéaire entre les variables X et Y . Lorsque $\rho(X, Y) = 0$, nous disons que les variables X, Y sont **non corrélées**.
- Une valeur $\rho(X, Y)$ positive indique que si X augmente, Y a tendance à augmenter également. Une valeur $\rho(X, Y)$ négative indique que si X augmente, Y a tendance à diminuer.

Théorème 89.3

Soient X, Y deux variables aléatoires réelles. Alors $X+Y$ est une variable aléatoire réelle et :

$$\text{Var}(X + Y) = \text{Var}(X) + 2\text{Cov}(X, Y) + \text{Var}(Y).$$

De même, $\text{Var}(X - Y) = \text{Var}(X) - 2\text{Cov}(X, Y) + \text{Var}(Y)$.

Application Le débruitage

Soient X, Y, W trois variables aléatoires telles que X représente un signal (par exemple la fréquence cardiaque d'un fœtus), W un bruit (par exemple la fréquence cardiaque de la mère), et $Y = X + W$ le signal mesuré, c'est-à-dire le signal bruité. Nous supposons que $\text{Var}(Y) \neq 0$, c'est-à-dire que Y n'est pas constant. Nous cherchons à retrouver le signal d'origine X , à

partir du signal bruité Y . Pour cela, nous supposons que le signal d'origine « ressemble » au signal bruité et qu'il est de la forme $\widehat{X} = aY + b$ où $a \neq 0$ et b sont des réels qui minimisent $\mathbb{E}((X - \widehat{X})^2) = \mathbb{E}((X - aY - b)^2)$. Montrer que $a = \frac{\text{Cov}(X, Y)}{\text{Var}(Y)}$ et $b = \mathbb{E}(X) - a\mathbb{E}(Y)$.

Nous devons minimiser la fonction définie sur $\mathbb{R}^2$ par

$$\begin{aligned} f(a, b) &= \mathbb{E}((X - aY - b)^2) \\ &= \text{Var}(X - aY - b) + (\mathbb{E}(X - aY - b))^2. \end{aligned}$$

Nous avons

$$\text{Var}(X - aY - b) = \text{Var}(X) + \text{Var}(aY + b) - 2\text{Cov}(X, aY + b) = \text{Var}(X) + a^2\text{Var}(Y) - 2a\text{Cov}(X, Y),$$

$$\text{et } \mathbb{E}(X - aY - b) = \mathbb{E}(X) - a\mathbb{E}(Y) - b.$$

Donc $f(a, b) = \text{Var}(X) + a^2\text{Var}(Y) - 2a\text{Cov}(X, Y) + (\mathbb{E}(X) - a\mathbb{E}(Y) - b)^2$, et la minimisation de f repose sur les deux conditions nécessaires :

$$\begin{cases} \frac{\partial f(a, b)}{\partial a} = 2a\text{Var}(Y) - 2\text{Cov}(X, Y) - 2\mathbb{E}(Y)(\mathbb{E}(X) - a\mathbb{E}(Y) - b) = 0 \\ \frac{\partial f(a, b)}{\partial b} = -2(\mathbb{E}(X) - a\mathbb{E}(Y) - b) = 0 \end{cases}$$

ce qui se résoud par

$$a = \frac{\text{Cov}(X, Y)}{\text{Var}(Y)} \text{ et } b = \mathbb{E}(X) - a\mathbb{E}(Y).$$

Pour s'entraîner (solutions p. 549)

89.1 Monsieur et Madame se donnent rendez-vous. Chacun arrive au hasard entre 12 h et 13 h, et attend son conjoint. Quel est le temps moyen de l'attente ? Nous noterons X et Y les instants d'arrivée (en heure) de Madame et Monsieur. Nous admettrons que la loi de (X, Y) est définie par la densité $f_{X,Y}$ telle que, pour tout $(x, y) \in \mathbb{R}^2$:

$$f_{X,Y}(x, y) = \mathbb{I}_{(x,y) \in [12;13]^2}.$$

89.2 Une firme qui produit des paquets de lessive a placé une figurine dans chaque paquet. Il y a 7 familles de figurines, toutes sont représentées en proportions égales dans l'ensemble des paquets de lessive. La directrice d'une blanchisserie achète n paquets de lessive. Quel est le nombre moyen de familles de figurines qui y seront représentées ?

89.3 Le but de cet exercice est de calculer des coefficients de corrélation entre deux variables aléatoires réelles X et Y dans deux cas particuliers : l'un où il n'existe aucune relation linéaire entre X et Y (nous trouverons $\rho(X, Y) = 0$), le second où la relation est linéaire (nous trouverons $\rho(X, Y) = \pm 1$).

1. Comme dans l'exercice détaillé de la Fiche 87, (X, Y) représente les coordonnées cartésiennes d'un point jeté au hasard dans le disque unité. Montrer que : $\rho(X, Y) = 0$.

2. Soit maintenant un couple de variables aléatoires (R, S) tel que $R = aS + b$ avec $a \neq 0$ et b des réels. Montrer que :

$$\rho(R, S) = +1 \text{ si } a > 0 \text{ et } \rho(R, S) = -1 \text{ si } a < 0.$$

89.4 Montrer toutes les propriétés du Théorème 89.1.

Variables aléatoires indépendantes

Nous avons déjà vu la notion d'indépendance entre deux événements, nous définissons maintenant la notion d'indépendance entre variables aléatoires. Les résultats présentés ici sont essentiels.



Fiche 74

1. Définition

Définition 90.1

Deux variables aléatoires réelles X et Y sont **indépendantes** si l'une des deux assertions équivalentes suivantes est vérifiée :

- quels que soient les deux ensembles A et B inclus dans $\mathbb{R}$, nous avons

$$\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A) \times \mathbb{P}(Y \in B),$$

- quels que soient x, y deux réels, nous avons

$$F_{X,Y}(x, y) = F_X(x) \times F_Y(y),$$

ce qui signifie $\mathbb{P}(X \leq x, Y \leq y) = \mathbb{P}(X \leq x) \times \mathbb{P}(Y \leq y)$.

Si les deux variables aléatoires sont discrètes, nous avons le théorème suivant.

Théorème 90.1

Deux variables aléatoires réelles **discrètes** X et Y sont **indépendantes** si et seulement si quels que soient les réels $x \in X(\Omega)$ et $y \in Y(\Omega)$, nous avons :

$$p_{X,Y}(x, y) = p_X(x) \times p_Y(y),$$

ce qui signifie $\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x) \times \mathbb{P}(Y = y)$.

Exemple

Reprenons l'exemple de la Fiche 86 où nous tirons deux cartes dans un jeu de 32 cartes, et où nous notons X et Y les nombres de rois et de dames tirés. Ces deux variables aléatoires ne sont pas indépendantes. Par exemple, $\mathbb{P}(X = 2, Y = 1) = 0 \neq \mathbb{P}(X = 2) \times \mathbb{P}(Y = 1)$.

Si les deux variables aléatoires sont à densité, nous avons le théorème suivant.

Théorème 90.2

Soient X, Y deux variables aléatoires réelles **à densité**, de densités respectives f_X, f_Y . Soit $f_{X,Y}$ la densité jointe du couple (X, Y) . Si quels que soient les réels x et y , nous avons :

$$f_{X,Y}(x, y) = f_X(x) \times f_Y(y),$$

alors X et Y sont **indépendantes**.

Exemple

Dans l'exemple de la Fiche 87, la densité jointe est $f_{X,Y}(x, y) = \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}}$. Les densités marginales sont $f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ et $f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-y^2/2}$. Nous avons donc $f_{X,Y}(x, y) = f_X(x) \times f_Y(y)$ pour tout couple de réels (x, y) . Les deux variables X et Y sont donc indépendantes. De plus, elles ont les mêmes densités marginales, donc elles ont la même loi, la loi normale centrée et réduite.


 Fiche 83

La Définition 90.1 de l'indépendance se généralise au cas de n variables aléatoires.

Définition 90.2

Les n variables aléatoires réelles $X_1, X_2, \dots, X_n$ sont **mutuellement indépendantes** ou **indépendantes** si quels que soient les ensembles $A_1, A_2, \dots, A_n$ inclus dans $\mathbb{R}$, nous avons

$$\mathbb{P}(X_1 \in A_1, X_2 \in A_2, \dots, X_n \in A_n) = \prod_{i=1}^n \mathbb{P}(X_i \in A_i).$$

Ceci est équivalent à :

$$\mathbb{P}(X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n) = \prod_{i=1}^n \mathbb{P}(X_i \leq x_i), \quad \text{pour tous réels } x_1, x_2, \dots, x_n.$$

2. Lois de Bernoulli et binomiale

 Fiche 81
Théorème 90.3

Soient $X_1, \dots, X_n$, n variables aléatoires réelles indépendantes, de loi de Bernoulli de paramètre p . Alors la variable aléatoire $X = \sum_{i=1}^n X_i$ suit une loi binomiale de paramètres $(n; p)$.

Exemple

Soient n expériences réalisées en laboratoire, indépendantes entre elles, toutes réalisées dans les mêmes conditions, et chacune menant à un succès avec probabilité p . Alors, le nombre de succès parmi les n expériences suit une loi binomiale de paramètres $(n; p)$. En effet, notons

$$X_i = \begin{cases} 1 & \text{si la } i^{\text{ème}} \text{ expérience aboutit à un succès,} \\ 0 & \text{sinon} \end{cases} \quad \text{et } X = \sum_{i=1}^n X_i.$$

X est donc le nombre de succès parmi les n expériences. Nous savons, par le Théorème 90.3, que X suit une loi binomiale de paramètres $(n; p)$.

3. Moments de variables aléatoires indépendantes**Théorème 90.4**

Soient X et Y deux variables aléatoires réelles indépendantes. Alors pour toutes fonctions g et h , nous avons

$$\mathbb{E}(g(X) \times h(Y)) = \mathbb{E}(g(X)) \times \mathbb{E}(h(Y)).$$

En particulier, si X et Y sont indépendantes alors $\mathbb{E}(XY) = \mathbb{E}(X) \times \mathbb{E}(Y)$.

Nous en déduisons les deux théorèmes suivants.

Théorème 90.5

Soient X et Y deux variables aléatoires réelles indépendantes. Alors nous avons

$$\text{Cov}(X, Y) = \rho(X, Y) = 0.$$



Autrement dit, **deux variables indépendantes sont non corrélées**. Mais attention, **la réciproque est fautive**, comme le montre l'Exercice 90.2.

Théorème 90.6

Si $X_1, \dots, X_n$ sont n variables aléatoires réelles indépendantes, alors nous avons

$$\mathbb{V}\text{ar}\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \mathbb{V}\text{ar}(X_i).$$

En particulier, si X et Y sont indépendantes, alors $\mathbb{V}\text{ar}(X + Y) = \mathbb{V}\text{ar}(X) + \mathbb{V}\text{ar}(Y)$.

Application

Roméo et Juliette

Roméo et Juliette ont rendez-vous, mais ils arrivent au hasard entre 20 h et 21 h, attendent 10 minutes et repartent s'ils n'ont pas vu arriver l'élue(e) de leur cœur. Quelle est la probabilité qu'ils se rencontrent ?

Notons X et Y les instants d'arrivée (en heures) de Juliette et Roméo. Tous deux arrivent au hasard entre 20 h et 21 h, donc X et Y suivent tous deux la loi uniforme sur $[20; 21]$:

$$\forall x \in \mathbb{R}, \quad f_X(x) = \mathbb{1}_{20 \leq x \leq 21}, \quad \text{et} \quad \forall y \in \mathbb{R}, \quad f_Y(y) = \mathbb{1}_{20 \leq y \leq 21},$$

où nous rappelons que $\mathbb{1}_{x \in A}$ est égal à 1 si $x \in A$, et 0 sinon.

De plus, X et Y sont indépendantes (car Roméo et Juliette arrivent au hasard), donc

$$\forall (x, y) \in \mathbb{R}^2, \quad f_{X,Y}(x, y) = f_X(x) \times f_Y(y) = \mathbb{1}_{20 \leq x \leq 21} \times \mathbb{1}_{20 \leq y \leq 21}.$$

Une fois arrivé(e), il (ou elle) attend 10 minutes (c'est-à-dire $\frac{1}{6}$ heure) puis repart. La probabilité qu'ils se rencontrent est donc, en notant $B = \{(x, y) \in [20; 21]^2 \text{ tels que } |x - y| \leq \frac{1}{6}\}$,

$$\mathbb{P}\left(|X - Y| \leq \frac{1}{6}\right) = \mathbb{P}((X, Y) \in B) = \iint_B f_{X,Y}(x, y) dx dy = \iint_B dx dy = \text{Aire de } B = \frac{11}{36}.$$

Le calcul de l'aire de B est expliqué par la Figure 90.1.

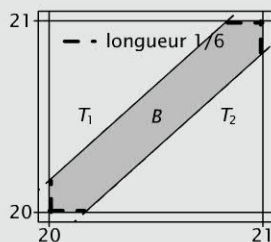


Figure 90.1 – L'aire de B est égale à l'aire du carré $[20; 21]^2$ moins l'aire des deux triangles T_1 et

$$T_2, \text{ soit Aire de } B = 1 - \left(1 - \frac{1}{6}\right)^2 = \frac{11}{36}.$$

Pour s'entraîner (solutions p. 549)

90.1 Soient $X_1, X_2, \dots, X_n$, n variables aléatoires réelles indépendantes, et de même loi. Notons F_X leur fonction de répartition commune : pour tout $i \in \{1; \dots; n\}$ et pour tout $x \in \mathbb{R}$, nous avons $F_X(x) = \mathbb{P}(X_i \leq x)$. Soient les deux variables aléatoires $Y = \min\{X_1; X_2; \dots; X_n\}$ et $Z = \max\{X_1; X_2; \dots; X_n\}$.

1. Calculer les fonctions de répartition des variables aléatoires Y et Z .

2. Soient n ampoules, dont les durées de vie suivent des lois exponentielles de même paramètre $\lambda > 0$. Soient Y le premier instant où l'une des ampoules grille, et Z le premier instant où toutes les ampoules sont grillées. Quelles sont les lois de Y et Z ?

90.2 Cet exercice montre que **deux variables aléatoires dépendantes peuvent être non corrélées**. Comme dans l'exercice détaillé de la Fiche 87, le couple aléatoire (X, Y) représente les coordonnées cartésiennes d'un point jeté au hasard dans le disque unité. Montrer que $\rho(X, Y) = 0$, mais que les deux variables X et Y ne sont pas indépendantes.

90.3

www

Soient X et Y les nombres de motos et de camions stationnés un certain jour entre 10 h et 12 h sur un immense parking. X et Y suivent des lois de Poisson de paramètres respectifs λ et μ . Quelle hypothèse vérifient X et Y ? Montrer que la loi du nombre de motos et de camions stationnés ce jour-là entre 10 h et 12 h est une loi de Poisson de paramètre $\lambda + \mu$. De même, soit Z le nombre de voitures stationnées ce jour-là entre 10 h et 12 h, sur ce parking. Z suit une loi de Poisson de paramètre γ . Quelle est la loi de $X + Y + Z$ le nombre de motos, camions et voitures stationnés ? Que représentent physiquement λ, μ et γ ?

90.4

www

Deux États notés État 1 et État 2 (n'ayant aucun contact entre eux) sont en concurrence pour réaliser une prouesse spatiale. À chaque tentative, ils ont respectivement les probabilités p_1 et p_2 de réussir. Les prouesses des deux États sont indépendantes. Quelle est la probabilité qu'un des États réussisse en moins de n tentatives ? Quelle est la probabilité que l'État 1 réussisse le premier ?

Fonction d'un couple aléatoire

Calcul de loi

Cette Fiche est l'analogue de la Fiche 80 dans le cas des couples aléatoires.

 Fiche 80

Exemple

Supposons qu'une machine produit des résistances dont la valeur en ohms est une variable aléatoire R . Une intensité aléatoire I traverse la résistance. Nous voulons savoir la loi de la tension $U = RI$. La différence entre cet exemple et celui de la Fiche 80 est qu'ici à la fois R et I sont aléatoires.

Dans cette Fiche, X et Y sont deux variables aléatoires réelles.

1. Méthodes de calcul de loi de $Z = g(X, Y)$

► Cas où $Z = g(X, Y)$ est discrète

Chercher la loi de Z revient à identifier $Z(\Omega)$ et déterminer pour tout $z \in Z(\Omega)$, la valeur de $\mathbb{P}(Z = z)$.

Il faut alors identifier l'événement $\{(X, Y) \in A\}$ équivalent à l'événement $\{Z = z\}$. Ensuite, nous calculons $\mathbb{P}(Z = z) = \mathbb{P}((X, Y) \in A)$.

Exemple

Soient les variables aléatoires X, Y représentant les nombres de deux-roues et de voitures sur un parking de cinéma un dimanche soir, et la variable aléatoire $Z = X + Y$. Supposons que le parking a au maximum 500 places, et que $X(\Omega) \times Y(\Omega) \subset \{0; \dots; 500\} \times \{0; \dots; 500\}$. Alors

$$Z(\Omega) = \{0; \dots; 500\} \text{ et pour tout } k \in Z(\Omega), \mathbb{P}(Z = k) = \sum_{j=0}^k \mathbb{P}(X = j, Y = k - j).$$

► Cas où $Z = g(X, Y)$ est continue

Soit (X, Y) un couple aléatoire continu dont nous connaissons la densité $f_{X,Y}$. Nous définissons la variable aléatoire $Z = g(X, Y)$, dont nous voulons connaître la loi.

Pour connaître la loi de Z , nous pouvons par exemple calculer sa fonction de répartition en tout point $z \in Z(\Omega)$: $F_Z(z) = \mathbb{P}(Z \leq z) = \iint_{\mathbb{R}^2} \mathbb{1}_{g(x,y) \leq z} f_{X,Y}(x, y) dx dy$.

2. Méthodes de calcul de loi de $(U, V) = h(X, Y)$

Soit (X, Y) un couple aléatoire. Nous définissons le couple aléatoire $(U, V) = h(X, Y)$, et nous cherchons à définir sa loi. Nous ne traiterons ici que le cas où les couples (X, Y) et (U, V) sont continus. **Nous supposons de plus que h est un $\mathcal{C}^1$ -difféomorphisme.** Rappelons que la loi d'un couple aléatoire (U, V) est définie par la donnée des probabilités $\mathbb{P}((U, V) \in B)$ pour tout ensemble B de $U(\Omega) \times V(\Omega)$:

$$\mathbb{P}((U, V) \in B) = \iint_B f_{U,V}(u, v) du dv.$$

 Fiche 60

Nous donnons ici la méthode pour obtenir $f_{U,V}$:

$$\begin{aligned}\mathbb{P}((U, V) \in B) &= \mathbb{P}(h(X, Y) \in B) \\ &= \mathbb{P}((X, Y) \in h^{-1}(B)) \\ &= \iint_{h^{-1}(B)} f_{X,Y}(x, y) dx dy \\ &= \iint_B f_{X,Y}(h^{-1}(u, v)) |J_{u,v}| du dv \text{ par le changement de variables } (x, y) = h^{-1}(u, v)\end{aligned}$$

où $J_{u,v} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix}$ est le Jacobien de la fonction h^{-1} , et $|J_{u,v}|$ la valeur absolue de ce Jacobien.

 Le changement de variables $(x, y) = h^{-1}(u, v)$ est légitime puisque h est un $\mathcal{C}^1$ -difféomorphisme.

Ceci étant valable pour tout ensemble B de $U(\Omega) \times V(\Omega)$, la densité $f_{U,V}$ du couple (U, V) est donc égale à, pour tout $(u, v) \in U(\Omega) \times V(\Omega)$:

$$f_{U,V}(u, v) = f_{X,Y}(h^{-1}(u, v)) |J_{u,v}| \mathbf{1}_{(u,v) \in U(\Omega) \times V(\Omega)}.$$

Application Le jeu de fléchettes

Soit un jeu de fléchettes dont la cible est placée sur l'origine du repère cartésien du plan. Soit (X, Y) la position en coordonnées cartésiennes de l'impact d'une fléchette. Supposons que X et Y suivent toutes les deux des lois normales de moyenne 0 et d'écart-type $\sigma > 0$ et sont indépendantes. Notons $R = \sqrt{X^2 + Y^2}$ la distance (en mètres) entre l'impact et la cible, et θ l'angle (en radians) entre le point d'impact et la cible, c'est-à-dire tel que $\tan \theta = \frac{Y}{X}$.

1. Quelle est la loi de (R, θ) ?
2. R et θ sont-elles indépendantes ? Quel est le coefficient de corrélation entre R et θ ?
3. Quelle est la probabilité que le point d'impact soit à moins d'un mètre de la cible et évite le demi-plan $y < 0$?
4. Dans quel rayon autour de la cible l'impact arrive-t-il dans plus de 90% des cas ? Que ce passe-t-il si σ diminue ?

Avant de commencer, donnons la loi du couple (X, Y) . Comme X et Y sont indépendantes, nous avons pour tout $(x, y) \in \mathbb{R}^2$, $f_{X,Y}(x, y) = f_X(x) \times f_Y(y)$. Comme X et Y suivent des lois normales, nous avons $f_{X,Y}(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$.

1. Posons

$$h : \begin{cases} \mathbb{R}^2 \longrightarrow \mathbb{R}^+ \times [0; 2\pi[\\ (x, y) \longmapsto (r, t). \end{cases}$$

La fonction h restreinte à $\mathbb{R}^2$ privé de la demi-droite $\{(x; 0) \text{ tel que } x \geq 0\}$ et à valeurs dans $\mathbb{R}_+^* \times]0; 2\pi[$ est un $\mathcal{C}^1$ -difféomorphisme, et

$$h^{-1} : \begin{cases} \mathbb{R}_+^* \times]0; 2\pi[\longrightarrow \mathbb{R}^2 \setminus \{(x; 0) \text{ tel que } x \geq 0\} \\ (r, t) \longmapsto (x, y) = (r \cos t, r \sin t). \end{cases}$$



Ce changement de variables est appelé **changement de variables polaire**. Le Jacobien de la fonction h^{-1} est égal à r .

Calculons, pour tout ensemble B de $\mathbb{R}_+^* \times]0; 2\pi[$, la probabilité $\mathbb{P}((R, \theta) \in B)$:

$$\begin{aligned}\mathbb{P}((R, \theta) \in B) &= \mathbb{P}((X, Y) \in h^{-1}(B)) \\ &= \iint_{h^{-1}(B)} f_{X,Y}(x, y) dx dy \\ &= \iint_B f_{X,Y}(r \cos t, r \sin t) r dr dt\end{aligned}$$

par le changement de variables polaire $(x, y) = (r \cos t, r \sin t)$.

Nous en déduisons que la densité de (R, θ) est $f_{R,\theta}(r, t) = r f_{X,Y}(r \cos t, r \sin t)$ pour tout $(r, t) \in \mathbb{R}_+^* \times]0; 2\pi[$. Donc, nous avons

$$f_{R,\theta}(r, t) = \frac{1}{2\pi\sigma^2} r e^{-\frac{r^2}{2\sigma^2}} \mathbf{1}_{r>0} \mathbf{1}_{0<t<2\pi}.$$

2. Calculons les lois marginales de (R, θ) . Nous avons, pour tout $(r, t) \in \mathbb{R}^2$:

$$\begin{aligned}f_R(r) &= \int_{\mathbb{R}} f_{R,\theta}(r, t) dt = \frac{r}{\sigma^2} e^{-\frac{r^2}{2\sigma^2}} \mathbf{1}_{r>0}, \\ f_\theta(t) &= \int_{\mathbb{R}} f_{R,\theta}(r, t) dr = \frac{1}{2\pi} \mathbf{1}_{0<t<2\pi}.\end{aligned}$$

Nous constatons que $f_{R,\theta}(r, t) = f_R(r) \times f_\theta(t)$ pour tout $(r, t) \in \mathbb{R}^2$, les deux variables R et θ sont donc indépendantes, et alors le coefficient de corrélation $\rho(R, \theta)$ est nul.

3. La probabilité que le point d'impact soit à moins d'un mètre de la cible et évite le demi-plan $y < 0$ est égale à

$$\mathbb{P}(R \leq 1, \theta \leq \pi) = \mathbb{P}(R \leq 1) \times \mathbb{P}(\theta \leq \pi)$$

par indépendance des variables R et θ . Nous avons donc

$$\mathbb{P}(R \leq 1, \theta \leq \pi) = \int_0^1 f_R(r) dr \times \int_0^\pi f_\theta(t) dt = \frac{1 - e^{-\frac{1}{2\sigma^2}}}{2}.$$

4. Cherchons la distance $r_{90\%}$ telle que $\mathbb{P}(R \leq r_{90\%}) \geq 0,9$. Nous devons résoudre

$$1 - e^{-\frac{r_{90\%}^2}{2\sigma^2}} \geq 0,9.$$

Nous obtenons $r_{90\%} \geq \sqrt{-2 \ln(0,1)} \sigma$. Si σ diminue, le rayon $r_{90\%}$ diminue : le tireur est plus précis.

Pour s'entraîner

(solutions p. 549)

91.1 Soit $f(x, y) = 2e^{-2x}e^{-y} \mathbb{I}_{(x, y) \in \mathbb{R}^+ \times \mathbb{R}^+}$ la densité d'un couple aléatoire (X, Y) . Calculer la fonction de répartition de $U = \frac{X}{Y}$. En déduire la densité de U .

91.2 1. Soit (X, Y) un couple de variables aléatoires à valeurs positives. Soit $f_{X, Y}$ la densité du couple (X, Y) . Calculer la loi de $\left(XY, \frac{Y}{X}\right)$. En déduire la loi de XY .

2. Application : dans l'exemple du début de cette Fiche, supposons que la résistance R est de loi uniforme sur $[20; 30]$ ohms, et que l'intensité I a pour densité $f_I(t) = \frac{4}{\sqrt{\pi}} t^{x-t^2} e^{-t^2} \mathbb{I}_{t > 0}$. Supposons que I et R sont indépendantes. Quelle est la loi de $U = RI$?

91.3 Soit un mobile de masse M et de vitesse V . Compte tenu des incertitudes sur les valeurs de M et V , nous supposons que M et V sont des variables aléatoires, de densités respectives f_M et f_V définies sur $\mathbb{R}_+^*$. Nous supposons que les variables M et V sont indépendantes. Notons $E = \frac{1}{2}MV^2$ l'énergie associée à ce mobile. Exprimer la densité de la variable E en fonction des densités f_M et f_V . Exprimer l'énergie moyenne du mobile en fonction de sa masse moyenne, sa vitesse moyenne et de la variance de sa vitesse.

91.4 Soit $S_1 = X_1$ l'instant de la première éruption volcanique enregistrée depuis l'installation d'un appareil de mesures d'activités sismiques. Soit S_2 le temps écoulé entre la première et la seconde éruptions. Supposons que S_1 et S_2 sont indépendantes, de même loi exponentielle de paramètre $\lambda > 0$. Quelle est la loi de $X_2 = S_1 + S_2$ l'instant de la deuxième éruption ?

91.5 Soit (X, Y) un couple aléatoire tel que X et Y sont indépendantes et de même loi. Notons f leur densité commune. Soient les variables aléatoires $U = \max(X, Y)$ et $V = \min(X, Y)$.

1. Quelle est la loi de (U, V) ?

2. U et V sont-elles indépendantes ?

91.6 Une grenouille pond N œufs, où N suit une loi de Poisson de paramètre $\lambda > 0$. Chaque œuf pondu a une probabilité $p \in]0; 1[$ d'éclore. Soit M le nombre d'œufs éclos.

1. Quelle est la loi de M ?

2. Quel est le nombre moyen d'œufs éclos ?

3. Quelle est la variance de M ?

Indication : Utiliser la formule des probabilités totales (voir théorème 74.3).

 Fiche 74

Loi des grands nombres

La loi des grands nombres est l'un des plus importants résultats des probabilités.

1. Inégalité de Bienaymé-Tchebychev

Théorème 92.1 Inégalité de Markov

Soit X une variable aléatoire à valeurs positives. Pour tout réel $a > 0$, nous avons $\mathbb{P}(X > a) \leq \frac{\mathbb{E}(X)}{a}$.

L'inégalité de Markov entraîne le théorème suivant :

Théorème 92.2 Inégalité de Bienaymé-Tchebychev

Soit X une variable aléatoire réelle d'espérance μ et de variance σ^2 . Alors pour tout réel $a > 0$, nous avons :

$$\mathbb{P}(|X - \mu| \geq a) \leq \frac{\sigma^2}{a^2}.$$

En pratique

- Plus la variance est faible, plus la probabilité que X s'écarte de sa moyenne est faible. Nous disons que la variance de X est un indicateur de la **dispersion** de X .
- L'inégalité de Bienaymé-Tchebychev s'écrit aussi, en prenant $a = b\sigma$ avec $b > 0$:

$$\mathbb{P}(X \in [\mu \pm b\sigma]) = \mathbb{P}(\mu - b\sigma \leq X \leq \mu + b\sigma) \geq 1 - \frac{1}{b^2}.$$

Par exemple, nous avons $\mathbb{P}(X \in [\mu \pm 2\sigma]) \geq \frac{3}{4}$, ou encore $\mathbb{P}(X \in [\mu \pm 3\sigma]) \geq \frac{8}{9}$.

- Les inégalités de Markov et Bienaymé-Tchebychev sont intéressantes si **seules la moyenne μ et la variance σ^2 de X sont connues**. Si la loi de X est connue, les valeurs des probabilités $\mathbb{P}(X > a)$ ou $\mathbb{P}(|X - \mu| \geq a)$ sont calculables, et les inégalités précédentes ne sont pas utiles. Par exemple, si X suit la loi $\mathcal{E}(1)$, alors $\mathbb{P}(X \geq 2) = e^{-2} \approx 0,135$, et l'inégalité de Markov donne : $\mathbb{P}(X \geq 2) \leq \frac{1}{2}$. Ainsi, les inégalités de Markov et Bienaymé-Tchebychev ne nécessitent pas de connaître la loi de X , mais les inégalités obtenues peuvent être grossières.

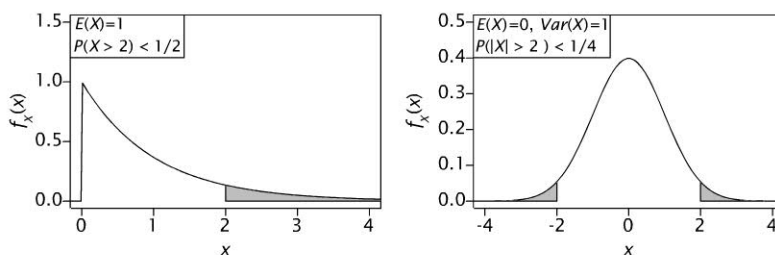


Illustration des inégalités de Markov et Bienaymé-Tchebychev : $\mathbb{P}(X \geq 2) \leq \frac{1}{2}$ (à gauche, via Markov),

$\mathbb{P}(|X| \geq 2) \leq \frac{1}{4}$ (à droite, via Bienaymé-Tchebychev).

Exemple

Supposons que la quantité hebdomadaire X produite par une usine est de 100 pièces en moyenne ($\mathbb{E}(X) = 100$), et que $\mathbb{V}\text{ar}(X) = 25$.

- La probabilité de produire plus de 150 pièces en une semaine vérifie : $\mathbb{P}(X > 150) \leq \frac{100}{150} = \frac{2}{3}$.
- La probabilité que la quantité de pièces produites soit comprise entre 100 ± 30 vérifie : $\mathbb{P}(70 \leq X \leq 130) \geq 1 - \frac{1}{6^2} = \frac{35}{36}$.

2. Loi des grands nombres

Commençons par définir les notions de convergence en probabilité et de convergence presque sûre.

Définition 92.1

Soient $(X_n)_{n \geq 0}$ une suite de variables aléatoires réelles et X une variable aléatoire réelle.

- La suite $(X_n)_{n \geq 0}$ **converge en probabilité** vers X si pour tout $a > 0$, $\lim_{n \rightarrow +\infty} \mathbb{P}(|X_n - X| \geq a) = 0$.
- La suite $(X_n)_{n \geq 0}$ **converge presque sûrement** vers X si $\mathbb{P}(\lim_{n \rightarrow +\infty} |X_n - X| = 0) = 1$.

Théorème 92.3 Loi faible des grands nombres

Soit $(X_n)_{n \geq 1}$ une suite de variables aléatoires réelles indépendantes et identiquement distribuées. Notons $\mu = \mathbb{E}(X_1)$ et $\sigma^2 = \mathbb{V}\text{ar}(X_1)$. Alors, nous avons pour tout réel $a > 0$, et tout entier $n > 0$:

$$\mathbb{P}\left(\left|\frac{1}{n} \sum_{i=1}^n X_i - \mu\right| \geq a\right) \leq \frac{\sigma^2}{na^2}.$$



- La **loi faible des grands nombres** est une application directe de l'inégalité de Bienaymé-Tchebychev.
- En particulier, la suite $\left(\frac{1}{n} \sum_{i=1}^n X_i\right)_{n \geq 1}$ converge en probabilité vers μ .

Théorème 92.4 Loi forte des grands nombres

Soit $(X_n)_{n \geq 1}$ une suite de variables aléatoires réelles indépendantes et identiquement distribuées. Notons $\mu = \mathbb{E}(X_1)$. Alors la suite $\left(\frac{1}{n} \sum_{i=1}^n X_i\right)_{n \geq 1}$ converge presque sûrement vers μ , i.e. avec probabilité 1, nous avons :

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{i=1}^n X_i = \mu.$$



Les théorèmes 92.3 et 92.4 sont **fondamentaux**. À retenir absolument !

L'application fondamentale de la loi des grands nombres est d'estimer une moyenne. Attention, la **moyenne empirique** $\frac{1}{n} \sum_{i=1}^n X_i$ est une variable aléatoire, tandis que la **moyenne théorique** μ est une constante réelle.

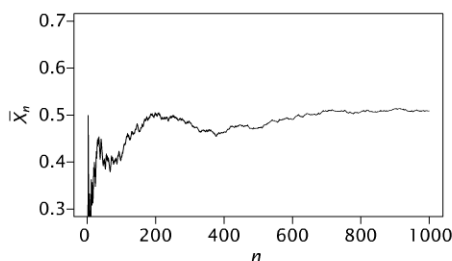


Illustration de la loi forte des grands nombres, avec $\mu = 0,5$ et une réalisation de $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$.

Application L'estimation d'une proportion

Une proportion $p \in [0; 1]$ de pièces sortant d'usine sont défectueuses. Nous réalisons n expériences qui consistent à tester si une pièce sortant de l'usine est défectueuse ou non. Posons $X_i = 1$ si la $i^{\text{ème}}$ pièce est défectueuse et $X_i = 0$ sinon. Soit $\alpha \in]0; 1[$, nous définissons un **intervalle de confiance de niveau $1 - \alpha$ de p** par un intervalle aléatoire $[A; B]$ formé de deux variables aléatoires réelles A et B telles que $\mathbb{P}(A \leq p \leq B) = 1 - \alpha$. Donner une estimation de p et un intervalle de confiance de niveau 95 %.

Les variables aléatoires X_i suivent toutes la loi de Bernoulli de paramètre p , et sont indépendantes (car nous supposons que les pièces sont indépendantes entre elles). Rappelons que $\mathbb{E}(X_i) = p$, et $\text{Var}(X_i) = p(1 - p)$.

La **loi forte des grands nombres** nous dit que, si n est assez grand,

$$\frac{1}{n} \sum_{i=1}^n X_i \approx \mathbb{E}(X_1) = p.$$

Une estimation de p est donc fournie par $\frac{1}{n} \sum_{i=1}^n X_i$. La **loi faible des grands nombres** avec

$a > 0$ que nous fixerons par la suite, nous donne $\mathbb{P}\left(\left|\frac{1}{n} \sum_{i=1}^n X_i - p\right| \leq a\right) \geq 1 - \frac{p(1-p)}{na^2}$.

Remarquons que nous avons : $\mathbb{P}\left(\left|\frac{1}{n} \sum_{i=1}^n X_i - p\right| \leq a\right) = \mathbb{P}\left(p \in \left[\frac{1}{n} \sum_{i=1}^n X_i \pm a\right]\right)$. L'intervalle

$\left[\frac{1}{n} \sum_{i=1}^n X_i \pm a\right]$ sera donc un intervalle de confiance de p , de niveau *supérieur ou égal* à $1 - \frac{p(1-p)}{na^2}$.

D'autre part, choisissons a tel que $1 - \frac{p(1-p)}{na^2} = 95\%$. Le problème est que p est inconnu.

Or, la fonction $x \mapsto x(1-x)$ définie sur $[0; 1]$ est majorée par $\frac{1}{4}$, donc

$$\mathbb{P}\left(p \in \left[\frac{1}{n} \sum_{i=1}^n X_i \pm a\right]\right) \geq 1 - \frac{1}{4na^2}.$$

Nous choisissons alors a tel que $1 - \frac{1}{4na^2} = 95\%$ ce qui donne $a = \sqrt{\frac{5}{n}}$. L'intervalle de confiance cherché est donc $\left[\frac{1}{n} \sum_{i=1}^n X_i \pm \sqrt{\frac{5}{n}} \right]$.



Le niveau de cet intervalle est *supérieur ou égal* à 95%. En effet, il est construit à partir de la loi faible des grands nombres qui fournit une inégalité, et non pas une égalité.

Pour s'entraîner (solutions p. 550)

92.1 La hauteur de pluie journalière dans une ville est de moyenne μ (en mm). Majorer la probabilité qu'il pleuve plus de 30 mm en 1 jour sur cette ville.

92.2 Soit X une variable aléatoire de moyenne $\mu \neq 0$ et écart-type $\sigma \neq 0$. Le rapport signal sur bruit de X est défini par $r = \frac{|\mu|}{\sigma}$. L'erreur relative du bruit sur le signal est définie par $E = \left| \frac{X - \mu}{\mu} \right|$. Montrer que pour tout $\epsilon > 0$, nous avons :

$$\mathbb{P}(E \leq \epsilon) \geq 1 - \frac{1}{r^2 \epsilon^2}.$$

92.3 

Soit une planche construite dans un matériau $\mathcal{M}$, et X la masse limite que peut supporter une planche. Supposons que l'écart-type de la variable X est égal à $\sqrt{\text{Var}(X)} = 80$ kg.

1. Supposons connue la masse limite moyenne : $\mathbb{E}(X) = 850$ kg.

a. Donner un intervalle de probabilité à 95 % de la masse limite X .

b. Soit une planche de même gabarit que la précédente. La masse qui a fait rompre la planche est 980 kg. Est-il possible que cette planche provienne du matériau $\mathcal{M}$?

2. Maintenant la masse limite moyenne $\mathbb{E}(X)$ est inconnue.

a. Soient n planches de même épaisseur, même longueur, même largeur, provenant

toutes du matériau $\mathcal{M}$. La masse de rupture de la planche i est noté X_i . Estimer $\mathbb{E}(X)$.

b. Quelle doit être la valeur de n pour que l'erreur entre l'estimation précédente et $\mathbb{E}(X)$ soit inférieure à 10^{-3} , avec une probabilité de 90 % ?

92.4 Nous lançons n points dans le carré $[0; 1]^2$, nous notons (X_i, Y_i) les coordonnées cartésiennes du $i^{\text{ème}}$ point, et $Z_i = \mathbb{1}_{X_i^2 + Y_i^2 \leq 1}$.

1. Quelle est la loi de Z_1 ? Calculer $\mathbb{E}(Z_1)$ et montrer que $\text{Var}(Z_1) \leq \frac{1}{4}$. Que dire de la suite $(Z_i)_{i \in \{1; \dots; n\}}$?

2. Posons $\bar{Z}_n = \frac{1}{n} \sum_{i=1}^n Z_i$. Montrer que pour tout $a > 0$,

$$\mathbb{P}(4\bar{Z}_n \in [\pi \pm 4a]) \geq 1 - \frac{1}{4na^2}.$$

3. Supposons que nous ne connaissons pas la valeur exacte de π . Après $n = 1000$ expériences, nous obtenons 0,786 pour $\bar{Z}_{1000}$. Quelle valeur approchée de π pouvons-nous proposer ? Déterminer un intervalle de confiance de niveau 95 % pour π (voir l'énoncé de l'exercice détaillé pour la définition d'intervalle de confiance).

4. Quelle taille n choisir pour que l'estimation soit proche de π à 10^{-2} près avec une probabilité de 95 % ?

Théorème de la limite centrale

Avec la loi des grands nombres, vue à la Fiche précédente, le Théorème de la limite centrale est l'un des plus importants résultats des probabilités.

1. Convergence en loi et Théorème de la limite centrale

Définition 93.1

Soit $(X_n)_{n \geq 1}$ une suite de variables aléatoires, et $(F_{X_n})_{n \geq 1}$ leurs fonctions de répartition respectives. Soit X une variable aléatoire et F_X sa fonction de répartition. La suite $(X_n)_{n \geq 1}$ **converge en loi** vers X si, en tout point $x \in \mathbb{R}$ où F_X est continue, nous avons $\lim_{n \rightarrow +\infty} F_{X_n}(x) = F_X(x)$.

Rappelons que si $X_1, \dots, X_n$ sont des variables aléatoires réelles indépendantes, de même loi, d'espérance μ et de variance $\sigma^2 > 0$, alors la **moyenne empirique** définie par $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$ vérifie $\mathbb{E}(\bar{X}_n) = \mu$ et $\text{Var}(\bar{X}_n) = \sigma^2/n$.

Théorème 93.1 Théorème de la limite centrale

Soit $(X_n)_{n \geq 1}$ une suite de variables aléatoires réelles indépendantes, de même loi, d'espérance μ et de variance $\sigma^2 > 0$. Alors la suite de variables aléatoires $\left(\frac{\bar{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}} \right)_{n \geq 1}$ converge en loi vers une loi normale centrée et réduite.



Théorème **fondamental**. À retenir absolument !



Point méthode

En pratique, nous écrivons que $\frac{\bar{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}} \sim \mathcal{N}(0; 1)$, ou $\bar{X}_n \sim \mathcal{N}\left(\mu; \frac{\sigma}{\sqrt{n}}\right)$ si $n \geq 30$.

Exemple

- Supposons que $X_1, \dots, X_n$ désignent $n \geq 30$ mesures indépendantes d'un même phénomène de moyenne μ et de variance σ^2 (par exemple un champ magnétique en un lieu donné), alors la mesure moyenne empirique définie par $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$ est approximativement distribuée selon une loi normale centrée en μ et de variance σ^2/n .
- Dans la Figure 93.1, nous avons choisi de représenter la densité de $\bar{X}_n$, pour différentes lois de X_1 . Nous avons $\mathbb{E}(\bar{X}_n) = \mu$ et $\text{Var}(\bar{X}_n) = \sigma^2/n$, donc $\bar{X}_n$ conserve la moyenne de la loi initiale et sa dispersion diminue lorsque n augmente. De plus la densité de $\bar{X}_n$ tend à s'approcher de la densité d'une loi normale.

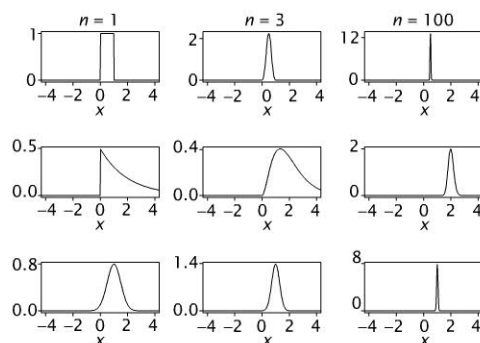


Figure 93.1 – Densité de $\bar{X}_n$, pour différents choix de $n = 1; 3$ et 100 et pour différentes lois : $X_1 \sim \mathcal{U}[0; 1]$ (première ligne), $X_1 \sim \mathcal{E}\left(\frac{1}{2}\right)$ (deuxième ligne) et $X_1 \sim \mathcal{N}\left(1; \frac{1}{2}\right)$ (troisième ligne).

2. Applications

Dans cette partie, Z est une variable aléatoire de loi normale centrée et réduite.

Supposons que nous étudions une variable aléatoire X , de moyenne μ et de variance σ^2 .

Soient une suite de variables aléatoires $X_1, \dots, X_n$ indépendantes et de même loi que

$$X, \text{ et } \bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i.$$

Si $n \geq 30$, nous savons par le Théorème de la limite centrale, que $\frac{\bar{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}}$ est approximativement distribuée selon une loi normale.

► Calcul de probabilité et de quantiles pour $\bar{X}_n$

Grâce au Théorème de la limite centrale, nous pouvons calculer approximativement $\mathbb{P}(\bar{X}_n \leq z)$ pour tout z réel. En effet, nous avons

$$\mathbb{P}(\bar{X}_n \leq z) = \mathbb{P}\left(\frac{\bar{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}} \leq \frac{z - \mu}{\frac{\sigma}{\sqrt{n}}}\right) \approx \mathbb{P}\left(Z \leq \frac{z - \mu}{\frac{\sigma}{\sqrt{n}}}\right),$$

et cette dernière valeur se lit dans la table de la loi normale centrée et réduite, donnée en annexe.

De même, étant donné un risque $\alpha \in [0; 1]$, nous pouvons chercher un quantile q_α d'ordre α pour $\bar{X}_n$. En effet, si z_α est un quantile d'ordre α de la loi normale centrée et réduite, alors $q_\alpha = \mu + \frac{\sigma}{\sqrt{n}} z_\alpha$, car

$$\mathbb{P}\left(\bar{X}_n \leq \mu + \frac{\sigma}{\sqrt{n}} z_\alpha\right) = \mathbb{P}\left(\frac{\bar{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}} \leq z_\alpha\right) \approx \mathbb{P}(Z \leq z_\alpha) = \alpha.$$

► Intervalle de confiance d'une moyenne

Supposons dans cette partie que μ est inconnue. Soit $\alpha \in [0; 1]$, et $z_{1-\frac{\alpha}{2}}$ le quantile d'ordre $1 - \frac{\alpha}{2}$ de la loi normale centrée et réduite.

Nous avons :

$$\begin{aligned}
 1 - \alpha &= \mathbb{P}(-z_{1-\frac{\alpha}{2}} \leq Z \leq z_{1-\frac{\alpha}{2}}) \\
 &\approx \mathbb{P}\left(-z_{1-\frac{\alpha}{2}} \leq \frac{\bar{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}} \leq z_{1-\frac{\alpha}{2}}\right) \\
 &\approx \mathbb{P}\left(\bar{X}_n - z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X}_n + z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right)
 \end{aligned}$$

Nous venons d'obtenir un **intervalle de confiance** de μ avec un **niveau de confiance** égal à $1 - \alpha$.

Définition 93.2

Soient $X_1, \dots, X_n$, $n \geq 30$ variables aléatoires indépendantes et de même loi, de moyenne μ inconnue, et de variance σ^2 connue. Nous définissons l'**intervalle de confiance de la moyenne μ , pour un niveau de confiance $1 - \alpha$** , avec $\alpha \in [0; 1]$, par l'intervalle aléatoire suivant :

$$\left[\bar{X}_n - z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}; \bar{X}_n + z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \right],$$

où $z_{1-\frac{\alpha}{2}}$ est le quantile d'ordre $1 - \frac{\alpha}{2}$ de la loi normale centrée et réduite.

Plus exactement, cet intervalle contient μ avec une probabilité $1 - \alpha$.

Application Le diabète

Lors d'une campagne d'information sur le diabète, une association mesure la glycémie à jeun de 100 personnes choisies au hasard dans la rue. Notons X_i la glycémie de la $i^{\text{ème}}$ personne interrogée. Nous calculons la moyenne empirique de la glycémie sur les 100 personnes par

$$\bar{X}_n = \frac{1}{100} \sum_{i=1}^{100} X_i.$$

Nous obtenons une moyenne empirique de 0,91 g/L. Nous connaissons

l'écart-type σ de la glycémie de la population : $\sigma = 0,06$ g/L.

1. Donner des intervalles de confiance à 90 % et 95 % de la glycémie moyenne de ce groupe.
2. Dans la suite, nous supposons que μ la glycémie moyenne dans la population est de 0,85 g/L. Cette valeur est-elle en accord avec les intervalles de confiance calculés précédemment ?
3. Pouvons nous résoudre l'exercice si le groupe de 100 personnes choisies au hasard est remplacé par un groupe de 30 personnes d'une même famille ?

Nous supposons que les variables aléatoires X_i sont **indépendantes**, car les personnes sont choisies au hasard. De plus, les variables X_i sont de **même loi** (toutes les personnes saines doivent avoir une glycémie de valeur similaire). Notons μ la glycémie moyenne de la population : $\mathbb{E}(X_i) = \mu$ pour tout $i = 1; \dots; n$ où $n = 100$.

1. Nous pouvons appliquer le **Théorème de la limite centrale**, et raisonner exactement comme dans la Section 2. Un intervalle de confiance de μ , de niveau $1 - \alpha$ est donné par

$$\left[\bar{X}_n - z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}; \bar{X}_n + z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \right]$$

avec $z_{1-\frac{\alpha}{2}} = 1,64$ ou $z_{1-\frac{\alpha}{2}} = 1,96$ respectivement pour les niveaux 90 % ou 95 %. Ces valeurs de $z_{1-\frac{\alpha}{2}}$ sont déterminées dans la table de la loi normale centrée et réduite, donnée en annexe. L'intervalle de confiance à 90 % de la glycémie moyenne est $[0,900; 0,920]$ g/L, tandis que l'intervalle à 95 % est $[0,898; 0,923]$ g/L.



Plus le niveau $1 - \alpha$ de l'intervalle de confiance est élevé, plus l'intervalle est large, ce qui est conforme à l'intuition.

2. La valeur théorique de la glycémie dans la population, 0,86 g/L n'est pas incluse dans les intervalles précédents. Nous concluons que le groupe interrogé a une glycémie supérieure à la moyenne.

3. Non : le nombre 30 est encore acceptable, mais les variables X_i ne sont plus indépendantes, car les personnes proviennent d'une même famille.

Pour s'entraîner (solutions p. 550)

93.1 Soit X_i le temps (en minutes) entre le $i^{\text{ème}}$ et le $i + 1^{\text{ème}}$ appels aux urgences d'un hôpital, et soit $S_n = \sum_{i=1}^n X_i$ l'instant du $n^{\text{ème}}$ appel aux urgences. Supposons que les variables aléatoires X_i sont indépendantes, et de loi exponentielle de paramètre $\lambda = 1$ (en minutes⁻¹). Quelle est, approximativement, la probabilité que les urgences reçoivent moins de 100 appels en 2 heures ?

93.2 Un ascenseur peut contenir une charge maximale de 600 kg. Supposons que n personnes montent dans l'ascenseur. Le poids de la $i^{\text{ème}}$ personne est modélisé par une variable aléatoire P_i de moyenne 60 kg, et d'écart-type 10 kg. Combien de personnes peuvent monter dans l'ascenseur pour que la charge maximale ne soit dépassée qu'avec une probabilité de 5 % ?

93.3 Le prix de l'action d'une société est donné pour le jour n par la variable aléatoire Y_n telle que $Y_n = Y_{n-1} + X_n$ où les variables aléatoires $X_1, \dots, X_n$ sont indépendantes, de même loi, d'espérance nulle et de variance $\sigma^2 = 1$. Supposons que l'action vaut aujourd'hui 1 euro. Que dire de la probabilité que l'action vaille plus de 2 euros dans un an ?

93.4

www

Un technicien d'une entreprise de télécommunication mesure le champ électrique émis par une antenne relais. Sachant les inévitables erreurs de mesure, il réalise plusieurs mesures autour de l'antenne. Notre technicien a des raisons de penser que les différentes valeurs mesurées sont des variables aléatoires indépendantes, de même loi, d'espérance commune μ et d'écart-type 2 V/m. Combien de mesures doit-il effectuer pour obtenir une approximation fiable à 99 % de μ avec une erreur inférieure à 0,5 V/m ?

93.5 Le calibre d'un œuf est fonction de son poids et détermine son prix de vente. Supposons que pour chaque œuf pondu, les probabilités que l'œuf soit respectivement petit, moyen, gros, ou extra gros sont : p_1, p_2, p_3 et p_4 telles que $\sum_{i=1}^4 p_i = 1$. Le prix de vente d'un œuf, qu'il soit petit, moyen, gros, ou extra gros est respectivement de v_1, v_2, v_3 et v_4 (en euros).

1. Quel est le prix moyen v de vente d'un œuf ?

2. Supposons qu'il n'y a pas de perte parmi les 500 œufs pondus du jour. Quelle est la probabilité que le fermier gagne plus de $500v$ euros avec la vente des 500 œufs du jour ?

Approximations d'une loi par une autre

L'intérêt pratique des résultats suivants est de remplacer certaines lois par d'autres, dans le but d'alléger les calculs, comme l'illustre l'Application de cette Fiche.

1. Approximation de la loi binomiale par la loi normale

Le théorème suivant est une conséquence directe du théorème central limite donné dans la Fiche 93.

Théorème 94.1

Soit $(Y_k)_{k \geq 1}$ une suite de variables aléatoires indépendantes, de loi de Bernoulli $\mathcal{B}(p)$ où $p \in]0; 1[$.

Soit $(X_n)_{n \geq 1}$ une suite de variables aléatoires, telle que pour tout $n \geq 1$, $X_n = \sum_{k=1}^n Y_k$.

Alors X_n suit la loi binomiale $\mathcal{B}(n; p)$. De plus, la suite $\left(\frac{X_n - np}{\sqrt{np(1-p)}} \right)_{n \geq 1}$ converge en loi vers une variable aléatoire de loi normale centrée et réduite $\mathcal{N}(0; 1)$.

En pratique

Lorsque n est grand, c'est-à-dire $n \geq 30$, et que p vérifie $np \geq 5$ et $n(1-p) \geq 5$, la loi binomiale $\mathcal{B}(n; p)$ peut être approchée par la loi normale $\mathcal{N}(np; \sqrt{np(1-p)})$.

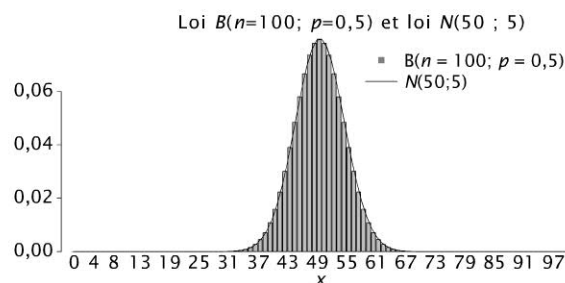


Figure 94.1 – Approximation de la loi binomiale par la loi normale.

2. Approximation de la loi de Poisson par la loi normale

Théorème 94.2

Soit $(X_n)_{n \geq 1}$ une suite de variables aléatoires, telle que pour tout $n \geq 1$, X_n suit la loi de Poisson de paramètre λ_n , avec $\lim_{n \rightarrow +\infty} \lambda_n = +\infty$. Alors, la suite $\left(\frac{X_n - \lambda_n}{\sqrt{\lambda_n}} \right)_{n \geq 1}$ converge en loi vers une variable aléatoire de loi normale centrée et réduite $\mathcal{N}(0; 1)$.

En pratique

Lorsque λ_n est grand, c'est-à-dire $\lambda_n \geq 20$, la loi de Poisson $\mathcal{P}(\lambda_n)$ peut être approchée par la loi normale $\mathcal{N}(\lambda_n; \sqrt{\lambda_n})$.

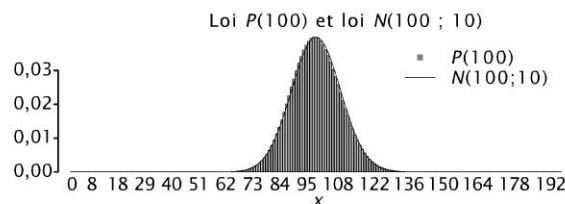


Figure 94.2 – Approximation de la loi de Poisson par la loi normale.

3. Approximation de la loi hypergéométrique par la loi binomiale

Soit E un ensemble contenant N éléments, dont une proportion p est de type A , et une proportion $1-p$ est de type B . Nous effectuons n tirages successifs et **sans remise** dans E . Soit X_N le nombre d'éléments de type A obtenus. Nous savons que la variable aléatoire X_N suit la loi hypergéométrique, notée $\mathcal{H}(N; p; n)$.

Fiche 81

Théorème 94.3

Notons X_N la variable aléatoire de loi $\mathcal{H}(N; p; n)$. Alors la suite $(X_N)_{N \geq 1}$ converge en loi vers une variable aléatoire de loi binomiale $\mathcal{B}(n; p)$.

Ainsi, lorsque N devient grand, en gardant fixes p la proportion d'éléments de type A , et n le nombre d'éléments prélevés, effectuer des **tirages sans remise** revient à effectuer des **tirages avec remise**.

En pratique

Lorsque $N \geq 10n$, la loi hypergéométrique $\mathcal{H}(N; p; n)$ peut être approchée par la loi binomiale $\mathcal{B}(n; p)$.

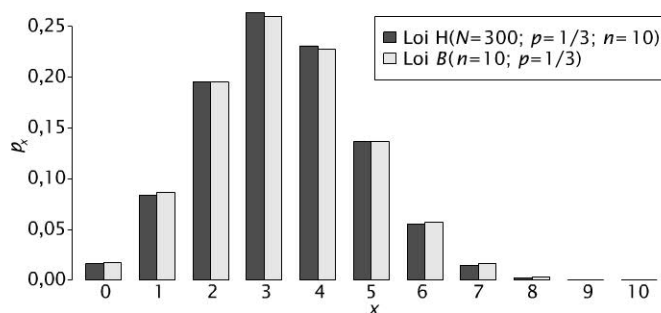


Figure 94.3 – Approximation de la loi hypergéométrique par la loi binomiale.

4. Approximation de la loi binomiale par la loi de Poisson

Théorème 94.4

Soit $\lambda > 0$ un réel fixé. Soit $(X_n)_{n \geq 1}$ une suite de variables aléatoires, telle que pour tout $n \geq 1$, X_n suit la loi binomiale $\mathcal{B}(n; p_n)$ où la suite des paramètres $(p_n)_{n \geq 1}$ est telle que

$\lim_{n \rightarrow +\infty} np_n = \lambda$. Alors la suite $(X_n)_{n \geq 1}$ converge en loi vers une variable aléatoire de loi de Poisson $\mathcal{P}(\lambda)$.

En pratique

Lorsque n est grand, c'est-à-dire $n \geq 30$, et que p est petit, c'est-à-dire $p \leq 0,1$, la loi binomiale $\mathcal{B}(n; p)$ peut être approchée par la loi de Poisson $\mathcal{P}(\lambda)$ où $\lambda = np$.

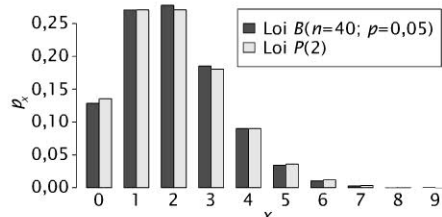


Figure 94.4 – Approximation de la loi binomiale par la loi de Poisson.

Application

Quelques pronostics sur une maladie rare

Une maladie assez rare affecte une personne sur 1000.

1. Sur une population de 5000 personnes choisies au hasard, à combien de personnes malades faut-il s'attendre ?
2. Quelle est la probabilité que 10 personnes soient malades ?
3. Quelle est la probabilité qu'au moins 3 personnes soient malades ?

Numérotions les personnes de 1 à $n = 5000$, et notons $X_i = 1$ si la $i^{\text{ème}}$ personne est malade, $X_i = 0$ sinon. Nous supposons que les variables X_i , pour $i = 1$ à n sont des variables aléatoires indépendantes (car les personnes sont choisies au hasard) et de même loi de Bernoulli de paramètre $p = 1/1000$.

1. Sur la population de n personnes, le nombre de malades est $X = \sum_{i=1}^n X_i$. La variable aléatoire X suit une loi binomiale de paramètres $(n; p)$. Nous savons, d'après la loi des grands nombres, que $\frac{X}{n} \approx \mathbb{E}(X_1) = p = \frac{1}{1000}$. Donc, nous nous attendons à $X \approx np = 5$ personnes malades.

2. Nous cherchons $\mathbb{P}(X = 10) = \binom{n}{10} p^{10} (1-p)^{n-10}$, d'après les formules de la loi binomiale.



Pour éviter de calculer ce terme (les factorielles cachées dans $\binom{n}{10}$ sont assez lourdes à calculer), nous allons approcher la loi de X par une loi de Poisson de paramètre $\lambda = np = 5$. Ceci est possible car $n \geq 30$, et $p \leq 0,1$.

Nous avons alors

$$\mathbb{P}(X = 10) \approx e^{-\lambda} \frac{\lambda^{10}}{10!} = 0,018.$$

3. Nous cherchons $\mathbb{P}(X \geq 3) = 1 - \mathbb{P}(X = 0) - \mathbb{P}(X = 1) - \mathbb{P}(X = 2)$. De même que dans la question précédente, nous utilisons l'approximation de la loi de X par la loi de Poisson de

Théo. 90.3,
Fiche 90

Fiche 92

paramètre λ . Nous avons alors

$$\mathbb{P}(X \geq 3) \approx 1 - e^{-\lambda} \left(\frac{\lambda^0}{0!} + \frac{\lambda^1}{1!} + \frac{\lambda^2}{2!} \right) = 0,875.$$

Pour s'entraîner (solutions p. 551)

94.1 Un élevage produit 500 œufs un jour donné. Supposons que, à chaque ponte, la probabilité que l'œuf soit cassé est $p = 0,05$.

1. Quelle est la loi de probabilité du nombre d'œufs cassés ce jour là ?

2. Quelle est la probabilité qu'il n'y ait pas plus de 20 œufs cassés ce jour là ?

94.2 La grippe saisonnière affecte 45 personnes sur 100. Quelle est la probabilité qu'au moins 29 258 022 personnes soient malades parmi la population française (qui compte 65 millions d'habitants) ?

94.3 Soit X le nombre de personnes passant par la gare de Paris Est lors d'un week-end de l'Ascension. Supposons que X suive une loi de Poisson de paramètre $\lambda = 3000$. Quelle est la probabilité qu'il y ait plus de 3100 personnes ?

94.4 Un lac contient $N = 100$ poissons dont une proportion $p = 0,1$ de brochets. Nous prélevons $n = 5$ poissons. Quelle est la probabilité d'obtenir $k = 2$ brochets ?

94.5  [www](#)

Suite à un incident de production dans une entreprise produisant des machines à laver, chaque machine sortie de l'atelier de production a une probabilité $p = 0,3$ d'avoir un défaut.

1. Sur un lot de 1000 machines à laver sorties d'usine, à combien de machines défectueuses faut-il s'attendre ?

2. Chaque machine défectueuse est renvoyée à l'usine, et cela coûte $v = 100$ euros à l'entreprise. Quel est le coût moyen de l'incident ?

3. Quelle est la probabilité que le coût de l'incident soit supérieur à 32 000 euros ?

Les tableaux ci-dessous donnent les moyennes et variances des lois discrètes et continues usuelles.

			cas discret	cas à densité
Espérance	$\mathbb{E}(X)$	$\sum_{x \in X(\Omega)} x \mathbb{P}(X = x)$	$\int_{\mathbb{R}} x f(x) dx$	
Espérance de $h(X)$	$\mathbb{E}(h(X))$	$\sum_{x \in X(\Omega)} h(x) \mathbb{P}(X = x)$	$\int_{\mathbb{R}} h(x) f(x) dx$	

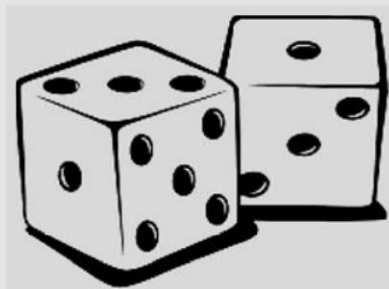
Lois discrètes	paramètre(s)	$X(\Omega)$	$\mathbb{P}(X = k)$	$\mathbb{E}(X)$	$\text{Var}(X)$
Uniforme $\mathcal{U}\{1; 2, \dots; n\}$	n	$\{1; 2, \dots; n\}$	$1/n$	$\frac{n+1}{2}$	$\frac{n^2-1}{12}$
Bernoulli $\mathcal{B}(p)$	$p \in]0; 1[$	$\{0; 1\}$	$\mathbb{P}(X = 1) = p$	p	$p(1-p)$
binomiale $\mathcal{B}(n; p)$	$n \in \mathbb{N}^*, p \in]0; 1[$	$\{0; 1, \dots; n\}$	$\binom{n}{k} p^k (1-p)^{n-k}$	np	$np(1-p)$
hypergéométrique	$N \in \mathbb{N}^*, p \in]0; 1[$	$\{\sup(0; n - (1-p)N); \dots$	$\frac{\binom{pN}{k} \times \binom{(1-p)N}{n-k}}{\binom{N}{n}}$	np	$np(1-p) \frac{N-n}{N-1}$
$\mathcal{H}(N; p; n)$	$n \in \{0; \dots; N\}$	$\dots; \inf(pN; n)\}$			
Poisson $\mathcal{P}(\lambda)$	$\lambda \in \mathbb{R}_+^*$	$\mathbb{N}$	$\exp(-\lambda) \frac{\lambda^k}{k!}$	λ	λ
géométrique $\mathcal{G}(p)$	$p \in]0; 1[$	$\mathbb{N}^*$	$p(1-p)^{k-1}$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
binomiale négative	$r \in \mathbb{N}^*, p \in]0; 1[$	$r; r+1; \dots$	$\binom{k-1}{r-1} p^r (1-p)^{k-r}$	$\frac{r}{p}$	$\frac{(1-p)r}{p^2}$

Lois à densité	paramètre(s)	support	densité $f(x)$	$\mathbb{E}(X)$	$\text{Var}(X)$
normale $\mathcal{N}(\mu; \sigma)$	$\mu \in \mathbb{R}, \sigma \in \mathbb{R}_+^*$	$\mathbb{R}$	$\frac{1}{\sqrt{2\pi}\sigma^2} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$	μ	σ^2
uniforme $\mathcal{U}[a; b]$	$a < b \in \mathbb{R}$	$[a; b]$	$\frac{1}{b-a} \mathbb{1}_{x \in [a; b]}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
exponentielle $\mathcal{E}(\lambda)$	$\lambda \in \mathbb{R}_+^*$	$\mathbb{R}^+$	$\lambda \exp(-\lambda x) \mathbb{1}_{x \geq 0}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Weibull $W(a; b)$	$a \in \mathbb{R}_+^*, b \in \mathbb{R}_+^*$	$\mathbb{R}^+$	$\frac{a}{b^a} x^{a-1} e^{-(x/b)^a} \mathbb{1}_{x \geq 0}$	$b \Gamma\left(1 + \frac{1}{a}\right)$	$b^2 \Gamma\left(1 + \frac{1}{a}\right) - \mathbb{E}(X)^2$
gamma $\Gamma(a; b)$	$a \in \mathbb{R}_+^*, b \in \mathbb{R}_+^*$	$\mathbb{R}^+$	$\frac{1}{b^a \Gamma(a)} x^{a-1} e^{-x/b} \mathbb{1}_{x \geq 0}$ avec $\Gamma(r) = \int_0^{+\infty} t^{r-1} e^{-t} dt$	ab	ab^2
beta $Beta(a; b)$	$a \in \mathbb{R}_+^*, b \in \mathbb{R}_+^*$	$[0; 1]$	$\frac{1}{B(a; b)} x^{a-1} (1-x)^{b-1} \mathbb{1}_{x \in [0; 1]}$ avec $B(a; b) = \int_0^1 x^{a-1} (1-x)^{b-1} dx$	$\frac{a}{a+b}$	$\frac{ab}{(a+b)^2(a+b+1)}$
Khi-deux $\chi^2(n)$	$n \in \mathbb{N}^*$	$\mathbb{R}^+$	$\frac{1}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})} x^{\frac{n}{2}-1} \exp\left(-\frac{x}{2}\right) \mathbb{1}_{x \geq 0}$	n	$2n$
Student $t(n)$	$n \in \mathbb{N}^*$	$\mathbb{R}$	$\frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})} \frac{1}{\sqrt{n\pi}} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}$	0 si $n > 1$	$\frac{n}{n-2}$ si $n > 2$
Fisher $F(n; m)$	$n \in \mathbb{N}^*, m \in \mathbb{N}^*$	$\mathbb{R}^+$	$\frac{\left(\frac{n}{m}\right)^{\frac{n}{2}}}{B\left(\frac{n}{2}, \frac{m}{2}\right)} \times \frac{x^{\frac{n}{2}-1}}{\left(1 + \frac{n}{m}x\right)^{\frac{n+m}{2}}} \mathbb{1}_{x \geq 0}$	$\frac{m}{m-2}$ si $m > 2$	$\frac{2m^2(n+m-2)}{n(m-2)^2(m-4)}$ si $m > 4$



Considérées aujourd'hui comme une branche à part entière des mathématiques, et ayant de nombreuses applications pratiques, les probabilités sont pourtant nées d'une activité d'apparence futile : les jeux de hasard.

Le mot « hasard » a d'ailleurs pour origine un mot arabe signifiant « dé ». Avant d'inventer le dé tel que nous le connaissons aujourd'hui, nos ancêtres pré-historiques jouaient avec un petit os, l'astragale, situé dans la cheville, ayant quatre faces régulières. Pourtant, si les jeux de hasard existent depuis les origines de l'humanité, il n'en est pas de même des Probabilités, qui elles, sont nées au XVII^e siècle.



La naissance des Probabilités

Les Probabilités sont aujourd'hui très répandues pour étudier des phénomènes aussi complexes que le climat, ou aussi simples que le résultat du tirage d'un dé. Leur langage est maintenant familier à tous les scientifiques confrontés à la modélisation de phénomènes aléatoires (météorologie, assurance, finance, télécommunication, médecine, etc.). Mais ce langage n'a pas toujours été familier.

En effet, ce n'est qu'au XVII^e siècle, suite à des réflexions soulevées par les jeux de hasard, que naissent les Probabilités. Blaise Pascal fut le premier à s'y intéresser, en cherchant à partager équitablement les mises des joueurs dans une partie de jeu inachevée, alors que l'un des joueurs a pris un avantage.



Blaise Pascal, 1623-1662

D'autres problèmes de cette époque, tous liés aux jeux de hasard, témoignent de l'effervescence liée à la naissance des Probabilités. Nous en citons deux ci-dessous : le paradoxe du Chevalier de Méré et le problème de Galilée, contemporains de Pascal.

Le paradoxe du Chevalier de Méré

Le Chevalier de Méré, qui était grand joueur, se posait la question suivante : « Est-il plus avantageux, lorsque l'on joue au dé, de parier sur l'apparition d'un 6 en lançant quatre fois le dé, ou bien de parier sur l'apparition d'un double 6, en lançant 24 fois deux dés ? ». Il avait remarqué que le premier pari était plus avantageux que le second, mais ne parvenait pas à comprendre pourquoi.

C'est Blaise Pascal qui trouva la réponse à la question du Chevalier de Méré. La probabilité d'apparition d'un 6 en lançant quatre fois le dé est $1 - \left(\frac{5}{6}\right)^4 \approx 0,5177 \dots$, tandis que la probabilité d'apparition d'un double 6 en 24 lancers de deux dés est $1 - \left(\frac{35}{36}\right)^{24} \approx 0,4914 \dots$. Le deuxième pari n'est donc pas avantageux !

Le problème de Galilée

Le Grand Duc de Toscane demanda un jour à Galilée : « Pourquoi, lorsque l'on jette trois dés, obtient-on plus souvent la somme 10 que la somme 9 ? Ces deux sommes sont pourtant obtenues toutes les deux avec six combinaisons de dés ! ». Galilée lui répondit par un texte rédigé entre 1613 et 1623. Les deux totaux sont en effet atteints avec six combinaisons de dés :

$$\begin{array}{ll}
 9 = 1 + 2 + 6 & 10 = 1 + 3 + 6 \\
 = 1 + 3 + 5 & = 1 + 4 + 5 \\
 = 1 + 4 + 4 & \text{et} \quad = 2 + 2 + 6 \\
 = 2 + 2 + 5 & = 2 + 3 + 5 \\
 = 2 + 3 + 4 & = 2 + 4 + 4 \\
 = 3 + 3 + 3 & = 3 + 3 + 4
 \end{array}$$

Pour une somme égale à 9, il faut que les tirages des trois dés donnent (1; 2; 6), (1; 3; 5), (1; 4; 4), (2; 2; 5), (2; 3; 4) ou (3; 3; 3). Pour une somme égale à 10, il faut que les tirages des trois dés donnent (1; 3; 6), (1; 4; 5), (2; 2; 6), (2; 3; 5), (2; 4; 4) ou (3; 3; 4).

Galilée remarqua que les combinaisons de dés ne sont pas toutes équivalentes, car les trois dés sont distincts. Par exemple, la combinaison (1; 2; 6) s'obtient ainsi : le 1 est donné par l'un des trois dés (trois choix), le 2 par l'un des deux dés restants (deux choix), et le 6 par le dernier dé restant (un seul choix), soit $3 \times 2 \times 1 = 3! = 6$ façons différentes d'obtenir la combinaison (1; 2; 6). De même pour les combinaisons (1; 3; 5) et (2; 3; 4). Ensuite, la combinaison (1; 4; 4) s'obtient de trois façons différentes : le 1 est donné par l'un des trois dés (trois choix), et les deux autres dés donnent obligatoirement 4 (un seul choix), soit 3 façons différentes d'obtenir la combinaison (1; 4; 4). De même pour la combinaison (2; 2; 5). Enfin, la combinaison (3; 3; 3) s'obtient d'une seule façon : les trois dés donnent 3. Ainsi, la somme des trois dés égale à 9 peut en définitive être obtenue de 25 façons différentes. Par le même raisonnement, Galilée montre que la somme des trois dés égale à 10 peut être obtenue de 27 façons différentes.

Ceci explique que la somme 10 survienne plus fréquemment que la somme 9.

À vous de réfléchir

Un pédiatre est appelé pour soigner les rhumes des cinq enfants de la famille Durand.

1. Il remarque quatre fillettes en train de jouer ensemble. Quelle est la probabilité que le cinquième enfant soit encore une fille ?
2. Il apprend ensuite que les quatre aînées sont des filles. Quelle est la probabilité que le petit dernier soit encore une fille ?

Statistique

Partie

4



Introduction

La statistique peut se diviser en deux branches : la **statistique descriptive** (Fiches 96 à 105) qui aborde les notions de population, d'échantillon, de variables statistiques, de graphiques, de caractéristiques de position, de dispersion et de forme. Puis vient la **statistique inférentielle** (Fiches 106 à 119) qui traite de l'estimation des caractéristiques d'une population à partir d'un échantillon.

Depuis plusieurs années, l'utilisation d'un logiciel est incontournable en statistique. Nous avons choisi le langage libre R. La fiche de présentation de R est disponible en bonus web. Le logiciel est téléchargeable à l'adresse suivante. Ce langage est compatible avec les systèmes d'exploitation Windows, Mac et Linux. Nous avons choisi ce langage pour deux raisons. La première est que le langage R est gratuit et ainsi, où que vous soyez dans le monde, vous pouvez le télécharger. La seconde est qu'il est de plus en plus enseigné à l'université ainsi que dans les grandes écoles. Ceci dit, les exercices proposés dans cette partie peuvent être résolus quel que soit le logiciel utilisé : Maple, Matlab, Minitab, SAS, SPSS, Statistica ...

Un focus consacré aux sondages illustre cette partie.

Les bonus web sur Dunod.com

 Retrouvez les corrigés détaillés d'une sélection d'exercices sur le site dunod.com ainsi que la fiche de présentation du logiciel R.

Concepts fondamentaux de la statistique

1. Population – Individus – Échantillon

Définition 96.1

L'ensemble sur lequel porte l'activité statistique s'appelle la **population**. Elle est généralement notée Ω pour rappeler la notation des probabilités mais par exemple dans la théorie des sondages elle est notée U , U comme « Univers ».

Fiche 72



- Le terme de « population » est employé aussi bien lorsqu'il s'agit d'un ensemble d'êtres humains que d'un ensemble d'objets inanimés.
- Il est recommandé de toujours bien définir la population étudiée.
- Le nombre d'éléments contenus dans Ω est généralement noté N , ce qui s'écrit également $\text{Card}(\Omega) = N$ ou encore $|\Omega| = N$.

Exemple

Dans les Fiches 100, 101 et 102, plusieurs caractères (définis par la suite) ont été recensés sur l'ensemble des pays européens. L'ensemble des pays européens constitue alors la population étudiée.

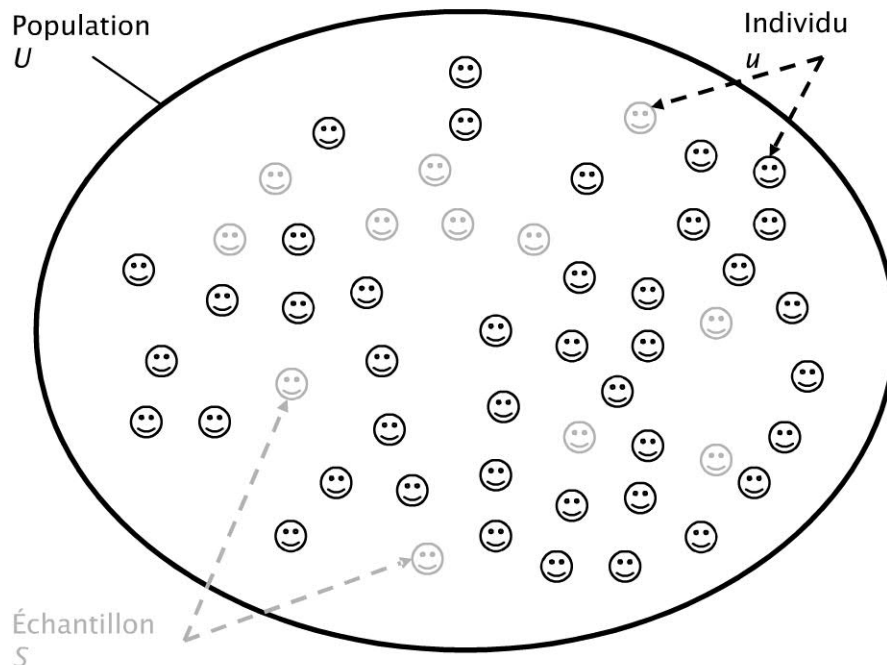


Figure 96.1 – Population, individus et échantillon.

Définition 96.2

Les éléments qui constituent la population sont appelés les **individus** ou encore les **unités statistiques**. Un individu est noté ω lorsque la population est notée Ω ou u lorsque la population est notée U .



Ces « individus » peuvent être de natures très diverses.

Exemples

Élèves en école d'ingénieurs, mois d'une année, pièces produites par une usine, résultats d'expériences répétées un certain nombre de fois, ...

Définition 96.3

Un **échantillon**, noté généralement $\mathcal{S}$, $\mathcal{S}$ comme « *sample* », échantillon en anglais, est une partie de la population prélevée soit de façon aléatoire, soit de façon déterministe.

Les trois notions de population, d'individus et d'échantillon sont représentées sur la Figure 96.1.

2. Observations – Caractères – Données brutes

Définition 96.4

Le statisticien fait des relevés sur les unités statistiques : ce sont les **observations**.

Définition 96.5

Les caractéristiques étudiées sur les individus d'une population sont appelées les **caractères**. Un **caractère** est une application χ d'un ensemble Ω fini de cardinal N (la population) dans un ensemble C (l'ensemble des valeurs possibles du caractère), qui associe à chaque individu ω de Ω la valeur $\chi(\omega)$ que prend ce caractère sur l'individu ω .

Nous considérons deux types de caractères :

1. Les **caractères quantitatifs** : leur détermination produit un nombre ou une suite de nombres. Nous distinguons
 - les **caractères simples** ou **univariés** : leur mesure sur un individu produit un seul nombre. Un caractère quantitatif simple est dit **continu** si l'ensemble C des valeurs possibles du caractère est un intervalle. Il est dit **discret** si l'ensemble C des valeurs possibles du caractère est discret.

Exemples

Taille, masse, salaire, température, ...

- les **caractères multiples** : leur mesure sur un individu produit une suite finie de nombres. L'ensemble de leurs valeurs est donc $\mathbb{R}^n$ ou une partie de $\mathbb{R}^n$.

Exemples

Relevé de notes d'un élève en école d'ingénieurs, fiche de salaire, ...

2. Les **caractères qualitatifs** simples ou multiples.

Exemples

Profession, adresse, sexe, numéro de téléphone, ...



La distinction entre les deux types de caractères est très importante. En effet, les méthodes d'analyse d'une population diffèrent suivant la nature du caractère étudié.

Exemple

Les représentations graphiques ne sont pas les mêmes.



Point méthode

Les caractères qualitatifs peuvent toujours être transformés en quantitatifs par codage. C'est ce qui se fait le plus généralement. Mais un tel codage est purement conventionnel et n'a pas vraiment un sens quantitatif.

Exemple

Nous ne pouvons pas calculer le sexe moyen.

Certains caractères qualitatifs s'expriment à l'aide de nombres.

Exemple

Un numéro de téléphone.

Mais ils n'ont pas non plus de sens quantitatif.

Exemple

Calculer un numéro de téléphone moyen n'est pas pertinent.

Définition 96.6

La suite des valeurs $\chi(\omega)$ prises par χ lorsque ω décrit toute la population Ω s'appelle les **données brutes**. C'est une suite finie $(X_1, \dots, X_N)$ d'éléments de l'ensemble C des valeurs possibles du caractère. Ces valeurs X_i ne sont pas nécessairement distinctes : un caractère peut prendre la même valeur sur deux individus différents.

3. Statistique descriptive – Statistique inférentielle

► Statistique ou statistiques ?

• Statistique

La **statistique** désigne à la fois un ensemble de données et l'ensemble des activités consistant à collecter ces données, à les traiter et à les interpréter.

• Statistiques

Les **statistiques**, l'ensemble des données numériques, interviennent pratiquement dans tous les domaines d'activité : gestion financière (états, banques, assurances, entreprises, ...), démographie, contrôles de qualité, études de marché, sciences expérimentales (biologie, psychologie ...) ...



Le mot « statistique » vient de l'allemand « *Statistik* », qui, au milieu du XVII^e siècle, désigne l'analyse de données utiles à l'État.

► Statistique descriptive

Le traitement des données, pour en dégager un certain nombre de renseignements qualitatifs ou quantitatifs généralement à des fins de comparaison, s'appelle la **statistique descriptive**. Elle ne s'applique que si les données ont été collectées sur la totalité des unités statistiques étudiées, c'est-à-dire la population en entier.

► Statistique inférentielle

Un autre but de la statistique consiste à extrapoler, à partir d'un échantillon de la population à étudier, le comportement de la population dans son ensemble. C'est la **statistique inférentielle** également appelée **statistique inductive**.

Exemples

Sondages, contrôle de qualité comportant un test destructif et plus généralement l'analyse des résultats de toute expérience pour laquelle il n'a pas été possible d'étudier la population en entier.

Application

Le tableau suivant représente la répartition des étudiants français ayant bénéficié du programme Erasmus de 2002/2003 à 2007/2008 par rapport aux pays d'accueil.

Pays \ Année	2002/ 2003	2003/ 2004	2004/ 2005	2005/ 2006	2006/ 2007	2007/ 2008
Espagne	4470	5115	5167	5481	5454	5281
Royaume-Uni	4705	4652	4564	4499	4673	4299
Allemagne	2808	2804	2863	2884	2800	2618
Italie	1415	1549	1571	1642	1638	1654

1. Quelle est la population étudiée ?
2. Quels sont les individus ?
3. Quels sont les caractères ?
4. Quelle est la nature des caractères ?

1. La population étudiée est l'ensemble des étudiants français ayant bénéficié du programme Erasmus de 2002/2003 à 2007/2008.
2. Les individus sont les étudiants français ayant bénéficié du programme Erasmus de 2002/2003 à 2007/2008.
3. Les deux caractères sont le « pays d'accueil » et l'« année universitaire pendant laquelle l'étudiant a bénéficié du programme Erasmus ».
4. Les deux caractères sont simples et qualitatifs.

Pour s'entraîner (solutions p. 551)

96.1 Pour chacune des séries statistiques présentées dans les Fiches 97 à 105, préciser les éléments suivants.

1. La population étudiée.
2. Les individus.
3. Le caractère ou les caractères.
4. La nature des caractères.

Statistique descriptive univariée

1. Modalités et classes

Le choix d'un caractère détermine le critère qui sert à comparer les individus entre eux.

Définition 97.1

Dans le cas d'un caractère qualitatif, cela revient à classer les individus en plusieurs groupes. Ils correspondent aux différentes valeurs que peut prendre le caractère qualitatif étudié. Ces dernières sont appelées **modalités**.

Exemples

Pour le caractère genre, il y a deux modalités. Pour la profession, la répartition peut se faire suivant plusieurs critères. En effet, elle peut, par exemple, se faire soit avec le nombre d'années d'étude après le baccalauréat, soit suivant les classes socio-professionnelles.



1. Le nombre de modalités peut-être fixé plus ou moins conventionnellement suivant la nature du caractère qualitatif étudié.
2. Les modalités du caractère étudié doivent être simultanément exhaustives (chaque individu appartient à une modalité car tous les cas ont été prévus) et incompatibles (un individu ne peut appartenir à deux modalités ou plus).

Exemple

Un élève en école d'ingénieur sera diplômé, ou il ne le sera pas.

Définition 97.2

Dans le cas d'un caractère quantitatif simple, discret ou continu, les individus sont souvent regroupés dans des **classes** c'est-à-dire au sein d'intervalles de valeurs que peut prendre le caractère. L'utilisation de classes est plus fréquente dans le cas d'un caractère quantitatif simple continu.

Exemple

Lors d'une étude du caractère quantitatif simple continu « taille » sur la population française, il est possible de regrouper les observations en classes de 10 cm de largeur de 0 à 230 cm. Ce découpage aboutirait aux classes suivantes : $]0 \text{ cm} ; 10 \text{ cm}]$, $]10 \text{ cm} ; 20 \text{ cm}]$, ..., $]220 \text{ cm} ; 230 \text{ cm}]$.

2. Tableaux statistiques

En fonction du contexte, il peut exister plusieurs possibilités pour représenter les données statistiques d'un échantillon à l'aide de tableaux. La solution qui peut être employée dans tous les cas, et qui l'est le plus souvent par les logiciels statistiques, est de présenter les tableaux de données sous la **forme empilée**. Dans cette forme, un **tableau statistique** est constitué :

- d'une ligne par individu,
- d'une colonne par caractère.

Concrètement, dans le cas de N observations d'un seul caractère nous pouvons avoir le **tableau de données brutes** suivant :

	X
Individu 1	X_1
Individu 2	X_2
...	...
Individu N	X_N

Dans le cas de N observations de deux caractères nous pouvons avoir le **tableau de données brutes** suivant :

	X	Y
Individu 1	X_1	Y_1
Individu 2	X_2	Y_2
...	...	...
Individu N	X_N	Y_N

Pour diminuer la place occupée par les tableaux, certaines personnes inversent malheureusement le rôle traditionnel des lignes et des colonnes dans les tableaux empilés. Si nous utilisons un logiciel pour réaliser nos analyses statistiques, il faudra penser à rétablir la convention précédente.

3. Série statistique

Définition 97.3

Les observations d'un ou plusieurs caractères récoltées sur toute la population forment une **série statistique**.

Les séries statistiques les plus élémentaires sont les séries à un seul caractère (**série statistique simple**) univarié (**série statistique simple univariée**) quantitatif ou qualitatif. Lorsque le caractère de la série statistique simple univariée est quantitatif discret (respectivement continu), la série est dite **série statistique simple discrète** (respectivement **continue**).

En général, ces séries sont représentées sous la forme d'un tableau statistique à une seule entrée puisqu'il n'y a qu'un seul caractère simple étudié. Parfois, la lecture de ces tableaux se révèle difficile. C'est pourquoi les représentations graphiques, qui sont présentées dans les Fiches 98, 99 et 102, sont utilisées très fréquemment pour faciliter l'analyse des informations qui sont contenues dans ces tableaux.

4. Description d'une série statistique quantitative

Soit X un caractère quantitatif simple discret ou continu. L'ensemble des valeurs atteintes par le caractère $X(\Omega) = \{X(\omega), \omega \in \Omega\} = \{X_1, \dots, X_N\}$ est un ensemble fini de valeurs distinctes $\{x_1, \dots, x_p\}$. Le fait que telle valeur soit relative à tel individu est un renseignement qui n'intéresse pas le statisticien. Seuls l'ensemble des valeurs atteintes et le nombre de fois que chacune d'elle est atteinte lui sont utiles. Nous supposons que les valeurs $x_1, \dots, x_p$ sont ordonnées : $x_1 < \dots < x_p$.

► Distribution statistique non groupée

Définition 97.4

Soit une série statistique simple, discrète ou continue, $(X_1, \dots, X_N)$ qui prend pour valeurs $x_1, \dots, x_p$ avec $x_1 < \dots < x_p$. Nous appelons :

1. **effectif de la valeur x_i** : le nombre n_i de fois que la valeur x_i est prise, c'est-à-dire le cardinal de l'ensemble $X^{-1}(x_i)$;
2. **effectif cumulé en x_i** : la somme $\sum_{j=1}^i n_j$;
3. **fréquence de la valeur x_i** : le rapport $f_i = \frac{n_i}{N}$ de l'effectif de x_i à l'effectif total N de la population, c'est-à-dire le cardinal de Ω ou encore la somme des n_i ;
4. **fréquence cumulée en x_i** : la somme $\sum_{j=1}^i f_j$.
5. La suite de couples $((x_i, n_i))_{i=1, \dots, p}$ ou $((x_i, f_i))_{i=1, \dots, p}$ est appelée **distribution statistique** (simple, discrète ou continue) **non groupée**.



1. Notons que, par définition, la somme des effectifs n_i est égale à N .
2. Notons que, par définition, la somme des fréquences f_i est égale à un.

Il est bien sûr évident que tous ces indicateurs peuvent figurer dans un même tableau statistique. Leur lecture est alors plus rapide car les informations sont concentrées dans un seul élément statistique.

► Distribution statistique groupée

Définition 97.5

Lorsque le caractère X quantitatif discret ou continu comprend un grand nombre de valeurs, il est préférable de regrouper ces valeurs en intervalles appelés **classes** pour rendre la statistique plus lisible. L'ensemble C des valeurs du caractère est alors partagé en classes $]a_i; a_{i+1}]$ avec $a_i < a_{i+1}$.

Définition 97.6

Soit une série statistique $(]a_i; a_{i+1}], n_i)_{i=1, \dots, p}$. Nous appelons :

1. **effectif de la classe $]a_i; a_{i+1}]$** : le nombre n_i de valeurs prises par le caractère dans l'intervalle $]a_i; a_{i+1}]$, c'est-à-dire $X^{-1}(]a_i; a_{i+1}])$;
2. **effectif cumulé en a_i** : le nombre de valeurs prises dans l'intervalle $] - \infty; a_i]$;
3. **fréquence de la classe $]a_i; a_{i+1}]$** : le rapport $f_i = \frac{n_i}{N}$;
4. **fréquence cumulée en a_i** : la somme $\sum_{j=1}^i f_j$.
5. La série statistique $(]a_i; a_{i+1}], n_i)_{i=1, \dots, p}$ ou $(]a_i; a_{i+1}], f_i)_{i=1, \dots, p}$ est appelée **distribution statistique** (simple, discrète ou continue) **groupée**.

5. Description d'une série statistique qualitative

Pour un caractère qualitatif simple, les notions d'effectif et de fréquence s'appliquent également en considérant les différentes modalités que peut prendre le caractère qualitatif. Si les modalités du caractère qualitatif peuvent être ordonnées, il est également naturel de se servir des effectifs cumulés et des fréquences cumulées. Il est aussi vivement recommandé de résumer toutes les informations concernant le caractère sous forme d'un tableau statistique.

Application

Les distributions statistiques

1. Distribution statistique non groupée

Population, en millions, des 27 États membres de l'Union Européenne donnée dans l'ordre du Tableau 99.1 de la Fiche 99. (82, 4 ; 8, 2 ; 10, 3 ; 7, 7 ; 0, 8 ; 5, 4 ; 40, 2 ; 1, 4 ; 5, 2 ; 66, 7 ; 10, 7 ; 10 ; 3, 9 ; 58 ; 2, 3 ; 3, 5 ; 0, 5 ; 0, 4 ; 16, 2 ; 38, 6 ; 10, 1 ; 22, 2 ; 59, 9 ; 10, 2 ; 5, 4 ; 1, 9 ; 8, 9).

- Calculer l'effectif et l'effectif cumulé de cette série statistique en $x_7 = 40, 2$.
- Calculer la fréquence et la fréquence cumulée de cette série statistique en $x_{10} = 66, 7$.

2. Distribution statistique groupée

Répartition des ingénieurs de moins de 65 ans selon l'âge en 2010.

(([20; 30], 187850), ([30; 35], 130050), ([35; 40], 115600), ([40; 45], 86700), ([45; 50], 65025), ([50; 55], 57800), ([55; 60], 43350), ([60; 65], 36125)).

- Comment interpréter les éléments du couple ([20; 30], 187850) ?
- Calculer l'effectif de la classe [45; 50] et l'effectif cumulé en 60 de cette série statistique.
- Calculer la fréquence de la classe [20; 30] et la fréquence cumulée en 50 de cette série statistique.

1.a. L'effectif en $x_7 = 40, 2$ est 1. L'effectif cumulé en $x_7 = 40, 2$ est 23.

b. La fréquence en $x_{10} = 66, 7$ est $1/27$. La fréquence cumulée en $x_{10} = 66, 7$ est $26/27$.

2.a. Le second élément, 187850, du couple ([20; 30], 187850) correspond à l'effectif de la classe qui apparaît comme le premier élément, [20; 30], du couple.

b. L'effectif de la classe [45; 50] est 65025. L'effectif cumulé en 60 est 686375.

c. La fréquence de la classe [20; 30] est $13/50$. La fréquence cumulée en 50 est $81/100$.

Pour s'entraîner (solutions p. 551)

97.1 Pour chacune des séries statistiques présentées dans les Fiches 98 à 105, répondre aux questions suivantes.

- Déterminer les effectifs et effectifs cumulés.
- Déterminer les fréquences et fréquences cumulées.

Représentations graphiques pour les séries statistiques quantitatives

1. Représentations graphiques pour les séries statistiques non groupées

Pour les séries statistiques non groupées, il existe deux types de représentation graphique :

1. le diagramme en bâtons ;
2. le polygone.

► Diagramme en bâtons des effectifs ou des fréquences

Définition 98.1

Le **diagramme en bâtons des effectifs**, voir Figure 98.1, (respectivement **des fréquences**, voir Figure 98.2) d'une distribution statistique non groupée est constitué d'une suite de segments verticaux d'abscisses x_i dont la longueur est proportionnelle à l'effectif (respectivement la fréquence) de x_i , pour i variant de 1 à p .

► Diagramme en bâtons des effectifs cumulés ou des fréquences cumulées

Définition 98.2

En remplaçant dans la définition précédente le mot effectifs (respectivement fréquences) par effectifs cumulés (respectivement fréquences cumulées), nous obtenons le **diagramme en bâtons des effectifs cumulés**, voir Figure 98.3, (respectivement **des fréquences cumulées**).

► Polygone des effectifs ou des fréquences

Définition 98.3

Le **polygone des effectifs** (respectivement **des fréquences**) d'une distribution statistique non groupée est obtenu à partir du diagramme en bâtons des effectifs (respectivement des fréquences, voir Figure 98.2), en joignant par un segment les sommets des bâtons.

► Polygone des effectifs cumulés ou des fréquences cumulées

Définition 98.4

En remplaçant dans la définition précédente le mot effectifs (respectivement fréquences) par effectifs cumulés (respectivement fréquences cumulées), nous obtenons le **polygone des effectifs cumulés**, voir Figure 98.3, (respectivement **des fréquences cumulées**).

2. Représentations graphiques pour les séries statistiques groupées

Pour les séries statistiques groupées, il existe deux types de représentation graphique :

1. l'histogramme ;
2. le polygone.

► Histogramme

Définition 98.5

L'**histogramme** est une représentation graphique d'une distribution statistique groupée. Il peut être utilisé pour estimer une densité ; il doit alors délimiter une aire totale égale à 1.

Premier cas : Les amplitudes des classes sont égales

Cet histogramme est constitué d'un ensemble de rectangles dont la largeur est égale à a , l'amplitude de la classe, et la hauteur égale à $K \times n_j$ où n_j est l'effectif de la classe, i varie de 1 à p , et K est un coefficient arbitraire (choix d'une échelle). L'aire totale sous l'histogramme est égale à $K \times N \times a$ où N est l'effectif total.

Second cas : Les amplitudes des classes sont inégales

Elles sont des multiples entiers $k_j \times a$ de l'amplitude a de l'une d'entre elles. Nous convenons, pour conserver le résultat précédent, de prendre pour hauteur du rectangle de la classe numéro j le quotient $(K \times n_j)/k_j$, par i variant de 1 à p , voir Figure 98.4.

► Polygone des effectifs ou des fréquences

Définition 98.6

Le **polygone des effectifs** ou **des fréquences**, voir Figure 98.5, d'une distribution statistique groupée est obtenu en joignant dans l'histogramme de cette distribution les milieux des côtés horizontaux supérieurs.

► Polygone des effectifs cumulés ou des fréquences cumulées

Définition 98.7

Le **polygone des effectifs cumulés** d'une distribution statistique groupée est la représentation graphique de la fonction définie sur chaque intervalle $]a_i; a_{i+1}]$, $1 \leq i \leq p$ par :

$$\forall x \in]a_i; a_{i+1}], \quad f(x) = \sum_{j=1}^{i-1} n_j + \frac{x - a_i}{a_{i+1} - a_i} n_i.$$

Définition 98.8

Le **polygone des fréquences cumulées**, voir Figure 98.6, d'une distribution statistique groupée est la représentation graphique de la fonction définie sur chaque

intervalle $]a_i; a_{i+1}]$, $1 \leq i \leq p$ par :

$$\forall x \in]a_i; a_{i+1}], \quad f(x) = \sum_{j=1}^{i-1} f_j + \frac{x - a_i}{a_{i+1} - a_i} f_i.$$

Application

1. Répartition des ménages selon le nombre de personnes en 2008

Voici la répartition des ménages, de moins de 6 personnes, en France métropolitaine selon le nombre de personnes en 2008, d'après le recensement de la population réalisé par l'INSEE :

Nombre de personnes	1	2	3	4	5	6
Nombre de ménages	8 914 849	8 776 420	3 911 419	3 303 676	1 240 991	468 123

Construire les représentations graphiques de cette série statistique à l'aide du diagramme en bâtons des effectifs, en superposant le diagramme en bâtons et le polygone des fréquences, en superposant le diagramme en bâtons et le polygone des effectifs cumulés.

2. Répartition des ingénieurs de moins de 65 ans selon l'âge en 2010

Voici la répartition des ingénieurs (personnes issues des écoles habilitées par la CTI) de moins de 65 ans selon l'âge, en France, d'après une étude de l'IESF réalisée en 2010. Le nombre total des ingénieurs diplômés de moins de 65 ans à la fin de l'année 2010 est 722 500.

Âge	20 à 30 ans	30 à 35 ans	35 à 40 ans	40 à 45 ans
Effectif	187 850	130 050	115 600	86 700

Âge	45 à 50 ans	50 à 55 ans	55 à 60 ans	60 à 65 ans
Effectif	65 025	57 800	43 350	36 125

Construire les représentations graphiques de cette série statistique à l'aide de l'histogramme des fréquences, en superposant le diagramme en bâtons et le polygone des fréquences et en superposant le diagramme en bâtons et le polygone des effectifs cumulés.

1. Les Figures 98.1, 98.2 et 98.3 correspondent aux représentations graphiques demandées.

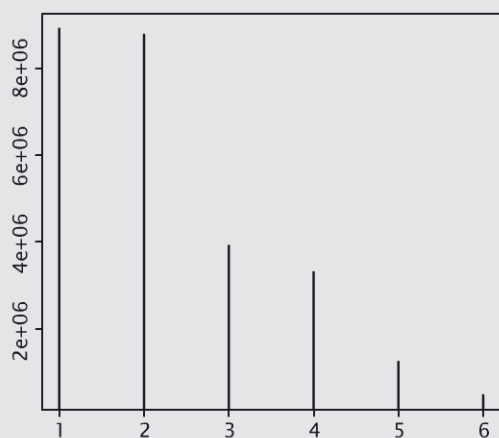


Figure 98.1

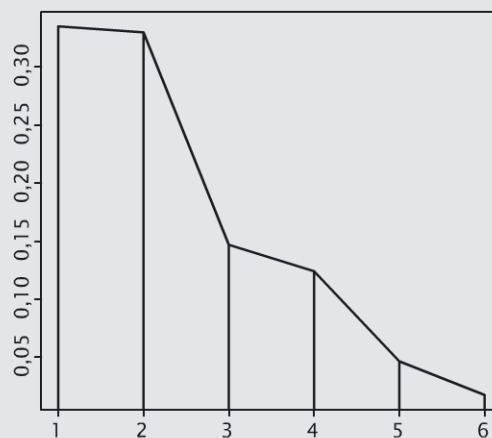


Figure 98.2

2. Les Figures 98.4, 98.5 et 98.6 correspondent aux représentations graphiques demandées.

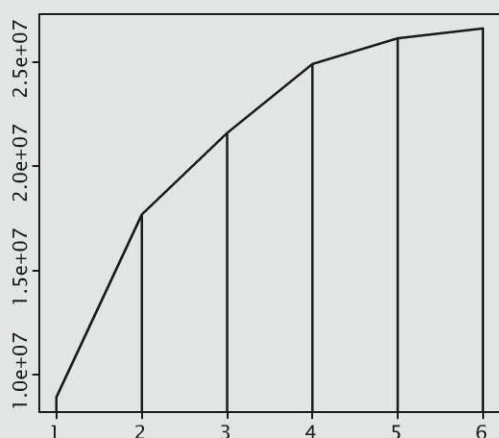


Figure 98.3

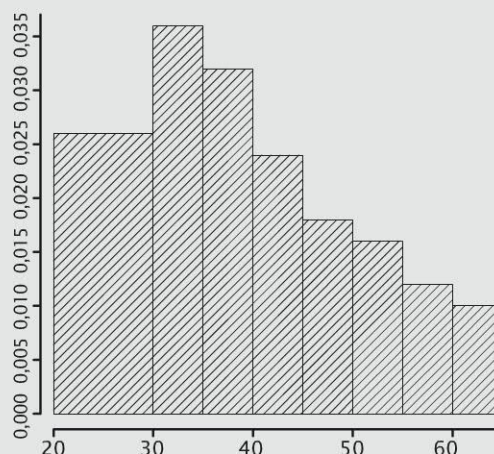


Figure 98.4

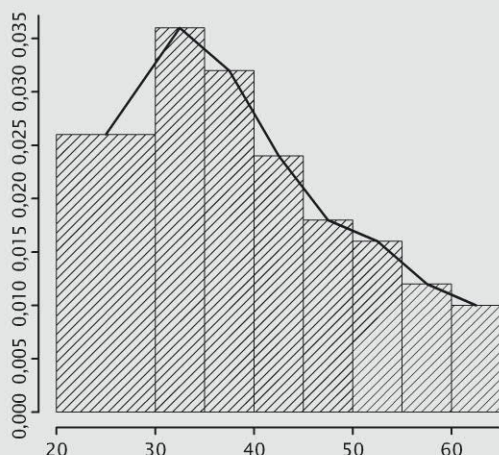


Figure 98.5

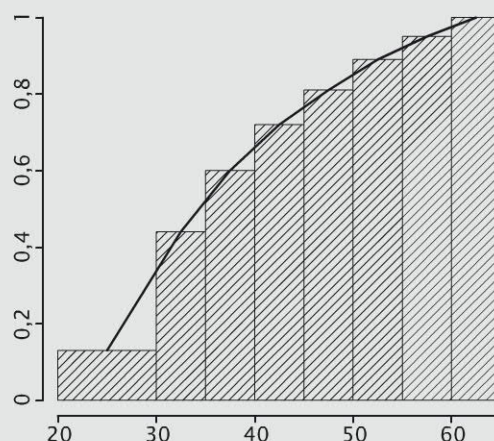


Figure 98.6

Pour s'entraîner (solutions p. 552)

98.1 Pour chacune des séries statistiques quantitatives présentées dans les Fiches 100 à 105, construire les représentations graphiques suivantes.

1. Si la distribution n'est pas groupée, le diagramme en bâtons des effectifs, le polygone des fréquences superposé au diagramme en

bâtons, le polygone des effectifs cumulés superposé au diagramme en bâtons.

2. Si la distribution est groupée, l'histogramme des fréquences, le polygone des fréquences superposé à l'histogramme et le polygone des effectifs cumulés superposé à l'histogramme.

Représentations graphiques pour les séries statistiques qualitatives

1. Diagramme à barres

Définition 99.1

Le **diagramme à barres verticales des effectifs** ou **des fréquences**, voir Figure 99.1, est une représentation graphique de la distribution, groupée ou non groupée, d'une série statistique qualitative simple par un ensemble de rectangles. L'axe des abscisses correspond aux différentes modalités du caractère de la série étudiée et l'axe des ordonnées aux effectifs ou fréquences associés.

Définition 99.2

Le **diagramme à barres horizontales des effectifs** ou **des fréquences**, voir Figure 99.4, est une représentation graphique de la distribution, groupée ou non groupée, d'une série statistique qualitative simple par un ensemble de rectangles. L'axe des ordonnées correspond aux différentes modalités du caractère de la série étudiée et l'axe des abscisses aux effectifs ou fréquences associés.

2. Diagramme à points

Définition 99.3

Le **diagramme à points vertical des effectifs** ou **des fréquences** est une représentation graphique de la distribution, groupée ou non groupée, d'une série statistique qualitative simple par un ensemble de points. L'axe des abscisses correspond aux différentes modalités du caractère de la série étudiée et l'axe des ordonnées aux effectifs ou fréquences associés.

Définition 99.4

Le **diagramme à points horizontal des effectifs** ou **des fréquences**, voir Figure 99.2, est une représentation graphique de la distribution, groupée ou non groupée, d'une série statistique qualitative simple par un ensemble de points. L'axe des ordonnées correspond aux différentes modalités du caractère de la série étudiée et l'axe des abscisses aux effectifs ou fréquences associés.

3. Diagramme circulaire

Définition 99.5

Le **diagramme circulaire** est une représentation graphique, voir Figure 99.3, de la distribution, groupée ou non groupée, d'une série statistique qualitative simple. À

chacune des modalités du caractère est associée une portion circulaire du diagramme proportionnelle à sa fréquence.



Les diagrammes circulaires ne sont pas des outils efficaces pour représenter les séries statistiques car l'œil humain distingue difficilement des variations d'angles ou de surfaces. De plus, il est toujours possible d'utiliser à leur place un diagramme à points, voir la section 2 de cette fiche.

Application

1. Les jeunes ingénieurs diplômés de 2010

Considérons la répartition par académie des diplômes délivrés en 2010 par les écoles d'ingénieurs.

- | | | |
|--------------------------|-----------------------|---------------------|
| • Aix-Marseille : 615 | • Lille : 1853 | • Reims : 504 |
| • Amiens : 1125 | • Limoges : 300 | • Rennes : 1664 |
| • Besançon : 803 | • Lyon : 2242 | • Rouen : 858 |
| • Bordeaux : 840 | • Montpellier : 712 | • Strasbourg : 669 |
| • Caen : 324 | • Nancy-Metz : 1500 | • Toulouse : 2096 |
| • Clermont-Ferrand : 589 | • Nantes : 1890 | • Paris : 2798 |
| • Dijon : 369 | • Nice : 361 | • Créteil : 1208 |
| • Grenoble : 1402 | • Orléans-Tours : 614 | • Versailles : 3114 |
| | • Poitiers : 462 | |

Construire les représentations graphiques de la distribution statistique non groupée de cette série statistique qualitative simple à l'aide d'un diagramme à barres verticales des effectifs, d'un diagramme à points horizontal des fréquences et d'un diagramme circulaire.

2. Répartition des ingénieurs de moins de 65 ans selon leur spécialisation en 2010

Considérons la répartition des ingénieurs de moins de 65 ans selon leur spécialisation en 2010.

- | | | |
|---------------------------------------|--------------------------------------|-------------------------------|
| • Agronomie, IAA, ... :
68550 | • STIC : 163570 | • Physique : 50270 |
| • Chimie : 49900 | • Génie civil :
50780 | • Économie, finance :
8930 |
| • Automatique, électricité :
61450 | • Mécanique, productique :
103230 | • Généraliste : 137890 |
| | | • Autre : 27830 |

Construire les représentations graphiques de la distribution statistique non groupée de cette série statistique qualitative simple à l'aide d'un diagramme à barres horizontales des fréquences, d'un diagramme à points horizontal des effectifs et d'un diagramme circulaire.

1. Les Figures 99.1, 99.2 et 99.3 correspondent aux représentations graphiques demandées.

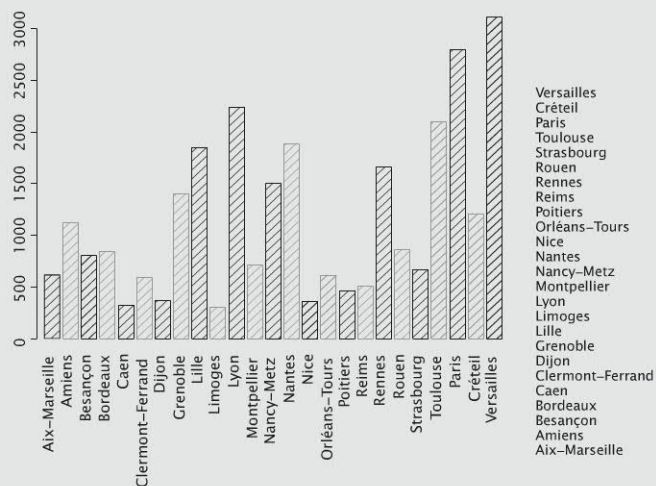


Figure 99.1

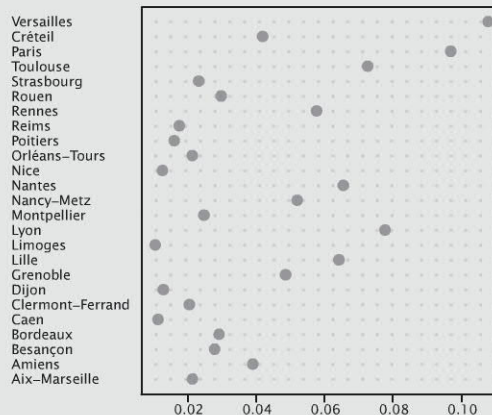


Figure 99.2

2. Les Figures 99.4, 99.5 et 99.6 correspondent aux représentations graphiques demandées.

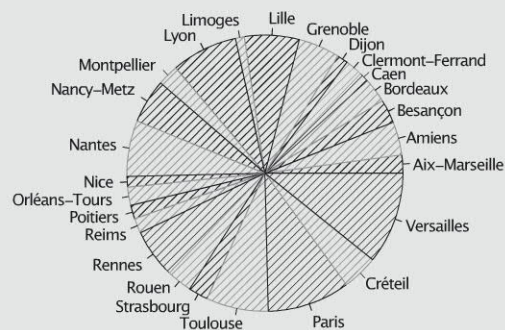


Figure 99.3

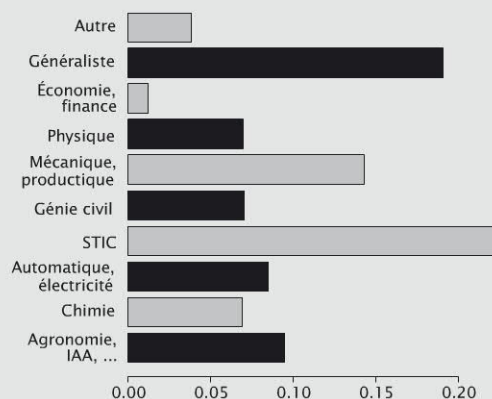


Figure 99.4

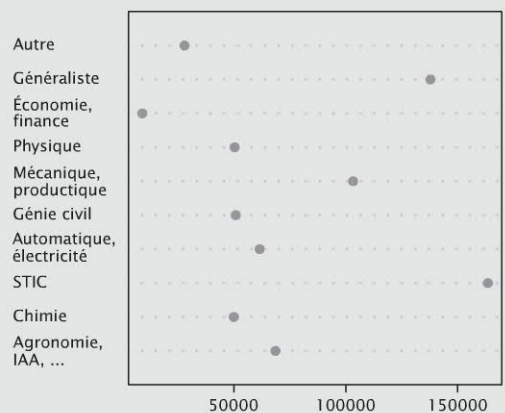


Figure 99.5

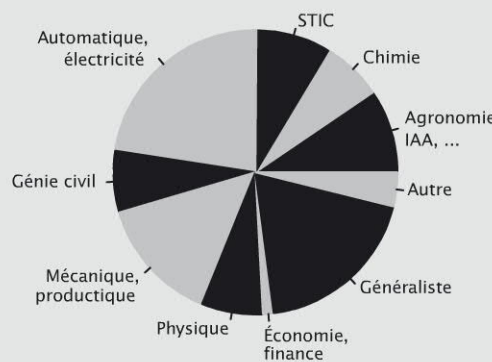


Figure 99.6

Pour s'entraîner (solutions p. 552)

99.1 Statistiques institutionnelles

Le Tableau 99.1, contient les valeurs de quelques statistiques institutionnelles pour les États membres de l'Union Européenne. Construire le diagramme à barres verticales

des fréquences, le diagramme à points horizontal des fréquences et le diagramme circulaire pour chacune de ces trois séries statistiques.

Tableau 99.1

États membres	Date d'entrée	Sièges au parlement	Voix au conseil
Allemagne	1957	99	29
Autriche	1995	18	10
Belgique	1957	24	12
Bulgarie	2007	—	—
Chypre	2004	6	4
Danemark	1973	14	7
Espagne	1986	54	27
Estonie	2004	6	4
Finlande	1995	14	7
France	1957	78	29
Grèce	1981	24	12
Hongrie	2004	24	12
Irlande	1973	13	7
Italie	1957	78	29
Lettonie	2004	9	4
Lituanie	2004	13	7
Luxembourg	1957	6	4
Malte	2004	5	3
Pays-Bas	1957	27	13
Pologne	2004	54	27
Portugal	1986	24	12
République Tchèque	2004	24	12
Roumanie	2007	—	—
Royaume-Uni	1973	78	29
Slovaquie	2004	14	7
Slovénie	2004	7	4
Suède	1995	19	10

Caractéristiques de position

1. Moyenne arithmétique

Définition 100.1

Notons N l'effectif de la population étudiée. La **moyenne arithmétique**, notée $\mu(X)$, d'une série statistique simple, discrète ou continue, dont la distribution statistique est *non groupée*, $(x_i, n_i)_{i=1, \dots, p}$, est le nombre réel défini par :

$$\mu(X) = \sum_{i=1}^p x_i f_i = \frac{1}{N} \sum_{i=1}^p x_i n_i.$$

Si la distribution est *groupée*, $([a_i; a_{i+1}], f_i)_{i=1, \dots, p}$, la **moyenne arithmétique** se calcule par :

$$\mu(X) = \sum_{i=1}^p \frac{a_i + a_{i+1}}{2} f_i.$$



Il existe d'autres moyennes comme la moyenne géométrique, la moyenne harmonique, la moyenne quadratique, la moyenne interquartile ou bien les moyennes tronquées.

2. Minimum et maximum

Définition 100.2

Le **minimum**, noté $\min(X)$, d'une série statistique est sa plus petite valeur.

Définition 100.3

Le **maximum**, noté $\max(X)$, d'une série statistique est sa plus grande valeur.

3. Médiane, classe médiane et statistique d'ordre

Considérons une distribution statistique simple, discrète ou continue, non groupée :

- **1^{er} cas** : N est un entier impair et de ce fait $N = 2m + 1$, où m est un entier.

Définition 100.4

La **médiane**, notée $Q_2(X)$ ou $Q_{0,5}(X)$, d'une série statistique est définie par : $Q_2(X) = x_{(m+1)}$ où $x_{(m+1)}$ signifie la $(m + 1)$ -ième valeur lorsque la série statistique est classée par ordre croissant.

Définition 100.5

$x_{(m+1)}$ est appelée la $(m + 1)$ -ième **statistique d'ordre de la série statistique**.

- **2^e cas** : N est un entier pair et de ce fait $N = 2m$, où m est un entier.

Définition 100.6

La **médiane**, notée $Q_2(X)$ ou $Q_{0,5}(X)$, d'une série statistique peut être prise égale à toute valeur supérieure ou égale à $x_{(m)}$ et inférieure ou égale à $x_{(m+1)}$. Toutefois, il est d'usage, bien qu'arbitraire, de prendre la moyenne de ces deux valeurs comme médiane : $Q_2(X) = (x_{(m)} + x_{(m+1)})/2$.

Considérons une distribution statistique simple, discrète ou continue, groupée :

Définition 100.7

Nous déterminons d'abord, à partir des fréquences cumulées, la classe contenant la médiane, *i.e.* la classe $]a_i; a_{i+1}]$ telle que $f_{i-1} < 0,5$ et $f_i \geq 0,5$, appelée **classe médiane**, puis la valeur de la médiane par interpolation linéaire dans cette classe, ce qui revient à supposer que les valeurs sont équiréparties au sein de cette classe.

4. Quartiles

Considérons une distribution statistique simple, discrète ou continue, non groupée :

- **1^{er} cas** : L'entier N n'est pas un multiple de 4. Soit m la partie entière de $N/4$.

Définition 100.8

Le **premier quartile**, noté $Q_1(X)$ ou $Q_{0,25}(X)$, d'une série statistique est défini par : $Q_1(X) = x_{(m+1)}$.

Soit m' la partie entière de $3N/4$.

Définition 100.9

Le **troisième quartile**, noté $Q_3(X)$ ou $Q_{0,75}(X)$, d'une série statistique est défini par : $Q_3(X) = x_{(m'+1)}$.

- **2^e cas** : L'entier N est un multiple de 4 et de ce fait $N = 4m$. Posons $m' = 3m = 3N/4$.

Définition 100.10

Le **premier quartile**, noté $Q_1(X)$ ou $Q_{0,25}(X)$, d'une série statistique peut être pris égal à toute valeur supérieure ou égale à $x_{(m)}$ et inférieure ou égale à $x_{(m+1)}$. Toutefois, il est d'usage, bien qu'arbitraire, de prendre la moyenne de ces deux valeurs comme premier quartile : $Q_1(X) = (x_{(m)} + x_{(m+1)})/2$.

Définition 100.11

Le **troisième quartile**, noté $Q_3(X)$ ou $Q_{0,75}(X)$, d'une série statistique peut être pris égal à toute valeur supérieure ou égale à $x_{(m')}$ et inférieure ou égale à $x_{(m'+1)}$. Toutefois, il est d'usage, bien qu'arbitraire, de prendre la moyenne de ces deux valeurs comme troisième quartile : $Q_3(X) = (x_{(m')} + x_{(m'+1)})/2$.

Considérons une distribution statistique simple, discrète ou continue, groupée :

Définition 100.12

Nous déterminons d'abord, à partir des fréquences cumulées, la classe contenant le premier (respectivement le troisième) quartile, *i.e.* la classe $]a_i; a_{i+1}]$ telle que $f_{i-1} < 1/4$ et $f_i \geq 1/4$ (resp. la classe $]a_i; a_{i+1}]$ telle que $f_{i-1} < 3/4$ et $f_i \geq 3/4$), appelée **classe du premier** (respectivement **du troisième**) **quartile**, puis la valeur du premier (respectivement du troisième) quartile par interpolation linéaire dans cette classe, ce qui revient à supposer que les valeurs sont équiréparties au sein de cette classe.



Le deuxième quartile est la médiane.

5. Mode et classe modale

Définition 100.13

Un **mode**, noté $Mo(X)$, d'une distribution statistique simple discrète *non groupée* est l'une des valeurs $x_1, x_2, \dots, x_p$ dont la fréquence est maximale.

Pour une distribution statistique simple *groupée*, une **classe modale**, $Mo(X)$, est une classe de densité, c'est-à-dire de rapport fréquence/longueur, maximale. Un mode est alors le centre de l'une de ces classes, ce qui revient à supposer que les valeurs sont équiréparties au sein de chaque classe.



Une distribution statistique simple, discrète ou continue, est **unimodale** si elle a un seul mode. Si elle en a plusieurs, elle est **plurimodale** (**bimodale**, **trimodale**, ...).

Application

1. Le temps de travail en Europe en 2010

Déterminer les caractéristiques de position de la variable, reproduite dans le Tableau 100.1, « Durée du travail par semaine pour les salariés ayant un emploi à plein temps en 2010 », à savoir :

- a. le minimum, le maximum et la moyenne arithmétique.
- b. La médiane, le premier quartile et le troisième quartile.

2. Répartition des ingénieurs de moins de 65 ans selon l'âge en 2010

Déterminer les caractéristiques de position de la variable introduite dans la seconde application de la Fiche 98, à savoir :

- a. le minimum, le maximum et la moyenne arithmétique.
- b. La médiane, le premier quartile et le troisième quartile.

1.a. Le minimum est $\min(X) = 37,70$ h ; le maximum est $\max(X) = 42,20$ h ; la moyenne arithmétique est $\mu(X) = 40,24$ h.

b. La médiane est $Q_{0,5}(X) = 40,50$ h ; le premier quartile est $Q_{0,25}(X) = 39,40$ h et le troisième quartile est $Q_{0,75}(X) = 41,10$ h.

2.a. Le minimum est $\min(X) = 20$ ans ; le maximum est $\max(X) = 65$ ans ; la moyenne arithmétique est $\mu(X) = 38,5$ ans.

b. Déterminons la classe médiane à partir de la distribution des fréquences cumulées : 0,26 ; 0,44 ; 0,60 ; 0,72 ; 0,81 ; 0,89 ; 0,95 ; 1,00. Par conséquent, la classe médiane est]35; 40]. Puis par interpolation linéaire nous obtenons $Q_{0,5}(X) = 35 + \frac{(40 - 35) \times (0,50 - 0,44)}{0,60 - 0,44} = 36,88$ ans.

À partir de la distribution des fréquences cumulées obtenue à la question précédente, nous déduisons que $Q_{0,25}(X)$ appartient à la classe]20; 30]. Puis par interpolation linéaire nous obtenons $Q_{0,25}(X) = 20 + \frac{(30 - 20) \times (0,25 - 0)}{0,26 - 0} = 29,62$ ans.

De même, $Q_{0,75}(X)$ appartient à la classe]45; 50]. Puis par interpolation linéaire nous obtenons $Q_{0,75}(X) = 45 + \frac{(50 - 45) \times (0,75 - 0,72)}{0,81 - 0,72} = 46,67$ ans.

Pour s'entraîner (solutions p. 552)

100.1 1. Calculer la valeur de la moyenne arithmétique de chacune des séries de données du Tableau 100.1 qui n'ont pas encore été étudiées.

2. Calculer les valeurs des trois quartiles de ces mêmes séries de données.

100.2 1. Calculer la valeur de la moyenne arithmétique de la série de données de la première application de la Fiche 98.

2. Calculer les valeurs des trois quartiles de cette même série de données.

Tableau 100.1 – Le temps de travail en Europe. Source : Eurostat.

États membres	Travailleurs à temps partiel en % de l'emploi total			Durée	
	Ensemble en (%)	Hommes en (%)	Femmes en (%)	Salariés (en heures)	Non-salariés (en heures)
Allemagne	26,2	9,7	45,5	40,6	50,8
Autriche	25,2	9,0	43,8	42,0	54,1
Belgique	24,0	9,0	42,3	39,2	55,9
Bulgarie	2,4	2,2	2,6	41,2	45,1
Chypre	9,3	6,5	12,7	41,1	48,6
Danemark	26,5	15,2	39,0	37,7	47,9
Espagne	13,3	5,4	23,2	40,4	48,5
Estonie	11,0	7,1	14,5	40,7	43,8
Finlande	14,6	10,0	19,6	39,1	48,3
France	17,8	6,7	30,1	39,4	53,4
Grèce	6,4	3,7	10,4	40,5	49,7
Hongrie	5,8	3,9	8,0	40,5	42,6
Irlande	22,4	11,8	34,7	38,4	47,9
Italie	15,0	5,5	29,0	39,0	45,5
Lettonie	9,7	7,8	11,4	40,4	42,6
Lituanie	8,1	6,7	9,3	39,6	40,9
Luxembourg	17,9	4,0	36,0	40,0	50,2
Malte	12,5	6,0	25,0	40,5	46,9
Pays-Bas	48,9	25,4	76,5	38,9	50,5
Pologne	8,3	5,7	11,5	40,9	47,3
Portugal	11,6	8,2	15,5	40,2	46,8
République Tchèque	5,9	2,9	9,9	41,2	48,6
Roumanie	11,0	10,6	11,4	41,1	40,0
Royaume-Uni	26,9	12,6	43,3	42,2	46,6
Slovaquie	3,9	2,8	5,4	40,6	45,6
Slovénie	11,4	8,6	14,7	41,1	46,9
Suède	26,4	14,0	40,4	39,9	48,8

Caractéristiques de dispersion

Les caractéristiques de position sont insuffisantes pour résumer de manière satisfaisante une série de données. En effet, deux séries de données ayant la même moyenne, ne se répartissent généralement pas de la même façon autour de celle-ci : elles sont plus ou moins étalées. Pour caractériser cette dernière propriété, il faut introduire les caractéristiques de dispersion.

1. Étendue

Définition 101.1

L'**étendue**, notée $e(X)$, est la différence entre la plus grande et la plus petite des valeurs prises par la série statistique simple, groupée ou non : $e(X) = \max(X) - \min(X)$.

2. Étendue interquartile et intervalle interquartile

Définition 101.2

L'**étendue interquartile**, notée $EIQ(X)$, est la différence entre le troisième quartile $Q_{0,75}(X)$ et le premier quartile $Q_{0,25}(X)$ de la série statistique simple, groupée ou non : $EIQ(X) = Q_{0,75}(X) - Q_{0,25}(X)$.

Définition 101.3

L'**intervalle interquartile** $[Q_{0,25}(X); Q_{0,75}(X)]$ contient 50 % des observations.

3. Variance

Définition 101.4

Soit N l'effectif de la population étudiée. Pour une distribution statistique simple, discrète ou continue, non groupée, la **variance**, notée $\text{Var}(X)$ ou $\sigma^2(X)$, est le nombre réel positif défini par :

$$\text{Var}(X) = \sigma^2(X) = \sum_{i=1}^p (x_i - \mu(X))^2 f_i = \frac{1}{N} \sum_{i=1}^p (x_i - \mu(X))^2 n_i.$$



En développant la formule ci-dessus, nous obtenons la **formule de Huygens** :

$$\text{Var}(X) = \sigma^2(X) = \sum_{i=1}^p x_i^2 f_i - \mu^2(X) = \mu(X^2) - \mu^2(X),$$

où $\mu(X^2)$ est la moyenne du carré des valeurs de la distribution et $\mu^2(X)$ est le carré de la moyenne de la distribution.

Cette dernière formule est souvent utilisée dans la pratique car elle simplifie les calculs.

Définition 101.5

Pour une distribution statistique simple groupée en classes d'amplitudes toutes égales à a , la **variance**, notée $\mathbb{V}\text{ar}(X)$ ou $\sigma^2(X)$, se calcule généralement en utilisant la **correction de Sheppard**. Elle est égale au nombre réel positif défini par :

$$\mathbb{V}\text{ar}(X) = \sigma^2(X) = \sum_{i=1}^p (x_i - \mu(X))^2 f_i - \frac{a^2}{12} = \frac{1}{N} \sum_{i=1}^p (x_i - \mu(X))^2 n_i - \frac{a^2}{12}.$$

Si les amplitudes des classes ne sont pas égales, alors la **variance**, notée $\mathbb{V}\text{ar}(X)$, se calcule généralement en utilisant une **généralisation de la correction de Sheppard** :

$$\mathbb{V}\text{ar}(X) = \sigma^2(X) = \frac{1}{N} \sum_{i=1}^p \left(\frac{a_i + a_{i+1}}{2} - \mu(X) \right)^2 n_i - \frac{1}{12N} \sum_{i=1}^p (a_{i+1} - a_i)^2 n_i.$$

4. Écart-type**Définition 101.6**

L'**écart-type** d'une distribution statistique simple, discrète ou continue, noté $\sigma(X)$, est égal à $\sqrt{\mathbb{V}\text{ar}(X)}$.



Il est souvent utilisé car il a la même unité que le caractère X .

5. Médiane des écarts absolus à la médiane**Définition 101.7**

La **médiane des écarts absolus à la médiane**, notée $\text{MAD}(X)$, d'une distribution statistique simple, discrète ou continue, non groupée est un nombre réel défini par :

$$\text{MAD}(X) = Q_{0,5} \left(\left(|x_i - Q_{0,5}(X)| \right)_{1 \leq i \leq n} \right).$$

6. Moment centré d'ordre r **Définition 101.8**

Le **moment centré d'ordre r** d'une distribution statistique, discrète ou continue, non groupée, noté ${}_{\mu}m_r(X)$, est égal à :

$${}_{\mu}m_r(X) = \sum_{i=1}^p (x_i - \mu(X))^r f_i.$$



1. Si $r = 2$, le moment centré d'ordre 2 n'est autre que la variance.
2. Tout comme les caractéristiques de dispersion permettent de représenter la distribution statistique plus fidèlement que les caractéristiques de position, les moments centrés d'ordre r supérieurs à deux améliorent cette représentation.



Fiche 102

7. Coefficient de variation

Définition 101.9

Le **coefficient de variation**, $CV(X)$, est défini pour un caractère X positif. Il est égal au rapport de l'écart-type $\sigma(X)$ avec la moyenne arithmétique $\mu(X)$.

Application

1. Le temps de travail en Europe en 2010

Pour la seconde application de la Fiche 100, établir les caractéristiques de dispersion suivantes.

- L'étendue, l'étendue interquartile.
- L'écart-type et la médiane des écarts absolus à la médiane.
- Le coefficient de variation exprimé en pourcents.

2. Répartition des ingénieurs de moins de 65 ans selon l'âge en 2010

Pour la seconde application de la Fiche 98, calculer les caractéristiques de dispersion suivantes.

- L'étendue, l'étendue interquartile.
- L'écart-type.
- Le coefficient de variation exprimé en pourcents.

1.a. L'étendue est 4,50 h ; l'étendue interquartile est $41,10 - 39,40 = 1,70$ h.

b. L'écart-type est 1,04 h ; la médiane des écarts absolus à la médiane est 0,60 h.

c. Le coefficient de variation exprimé en pourcents est 2,58 %.

2.a. L'étendue est 45 ans, l'étendue interquartile est 17,05 ans.

b. L'écart-type est 11,37 ans.

c. Le coefficient de variation exprimé en pourcents est 29,54 %.

Pour s'entraîner (solutions p. 552)

101.1

1. Calculer la valeur de l'écart-type de la série de données de la première application intitulée « Répartition des ménages selon le nombre de personnes en 2008 », de la Fiche 98.

2. Calculer la valeur de la médiane des écarts absolus à la médiane de cette même série de données.

3. Calculer la valeur de l'étendue interquartile de cette même série de données.

4. Calculer la valeur du coefficient de variation exprimé en pourcents de cette même série de données.

101.2

1. Calculer la valeur de l'écart-type de chacune des séries de données reproduite dans le

Tableau 100.1 à l'exclusion de celle déjà étudiée à l'application précédente.

2. Calculer la valeur de la médiane des écarts absolus à la médiane de ces mêmes séries de données.

3. Calculer la valeur de l'étendue interquartile de ces mêmes séries de données.

4. Calculer la valeur du coefficient de variation exprimé en pourcents de ces mêmes séries de données.

101.3 Nous avons regroupé en classes, de même amplitude, les caractères qui ont été introduites avec le nouveau jeu de données sur le Travail en Europe à la Fiche 100 et que vous avez déjà étudiés dans cette fiche. Les distributions groupées obtenues ont été reproduites ci-dessous.

1. À partir de la distribution groupée, calculer la valeur de la moyenne arithmétique de chacune des séries de données.

2. À partir de la distribution groupée, calculer les valeurs des trois quartiles de cette même série de données.

3. À partir de la distribution groupée, calculer la valeur de l'écart-type de chacune des séries de données.

4. À partir de la distribution groupée, calculer la valeur de l'étendue interquartile de ces mêmes séries de données.

5. À partir de la distribution groupée, calculer la valeur du coefficient de variation exprimé en pourcents de ces mêmes séries de données.

6. Comparer ces valeurs avec celles que vous avez obtenues à l'exercice précédent à partir de la distribution non groupée.

7. Conclure.

Fiche 100

Temps partiels ensemble (% emplois totaux)

[0;10]	[10;20]	[20;30]	[30;40]	[40;50]
9	10	7	0	1

Temps partiels hommes (% emplois totaux)

[0;10]	[10;20]	[20;30]
21	5	1

Temps partiels femmes (% emplois totaux)

[0;10]	[10;20]	[20;30]	[30;40]	[40;50]	[50;60]	[60;70]	[70;80]
5	9	3	4	5	0	0	1

Temps de travail salariés (heures/semaines)

[35;40]	[40;45]
10	17

Temps de travail non-salariés (heures/semaines)

[35;40]	[40;45]	[45;50]	[50;55]	[55;60]
1	4	16	5	1

Source : Eurostat.

Boîte à moustaches et caractéristiques de forme

Pour tracer une boîte à moustaches, il nous faut avoir acquis les notions de caractéristiques de position et de dispersion.

 Fiches 100
et 101

1. Définition d'une boîte à moustaches

La **boîte à moustaches** est un graphique sur lequel sont représentées de nombreuses caractéristiques de position et de dispersion. En fait, la boîte à moustaches est un moyen rapide de figurer le profil essentiel d'une série statistique X quantitative simple, discrète ou continue.



La boîte à moustaches peut être aussi appelée boîte de distribution, boîte de dispersion ou boxplot.

Une question à se poser est : « comment cette boîte à moustaches est-elle construite ? » Il en existe plusieurs versions, mais une telle boîte comprend le plus souvent :

- une échelle des valeurs présente sur l'axe vertical,
- la valeur du premier quartile, $Q_{0,25}(X)$, qui est le trait inférieur de la boîte,
- la valeur du troisième quartile, $Q_{0,75}(X)$, qui est le trait supérieur de la boîte,
- la valeur de la médiane, $Q_{0,5}(X)$, représentée par un trait horizontal au sein de la boîte,
- les moustaches inférieure et supérieure représentées par des pointillés ou des traits pleins verticaux de chaque côté de la boîte et qui se terminent par des traits horizontaux,
- les valeurs atypiques représentées chacune soit par un cercle soit par une étoile.



Parfois d'autres caractéristiques sont représentées comme la moyenne.

2. Construction d'une boîte à moustaches

La **boîte à moustaches** se représente en suivant les instructions suivantes :

1. Nous traçons un rectangle de largeur fixée à priori et de longueur, l'étendue interquartile, à savoir $EQ(X) = Q_{0,75}(X) - Q_{0,25}(X)$.
2. Nous y situons la médiane par un segment positionné à la valeur $Q_{0,5}(X)$, par rapport à $Q_{0,75}(X)$ et $Q_{0,25}(X)$. Parfois nous y ajoutons aussi la moyenne.
3. Nous calculons $(Q_{0,75}(X) + 1,5 \times EQ(X))$ et nous cherchons la dernière observation x_h en deçà de la limite $(Q_{0,75}(X) + 1,5 \times EQ(X))$, soit $x_h = \max \{x_i | x_i \leq Q_{0,75}(X) + 1,5 \times EQ(X)\}$.

 Fiche 101

 Fiche 100

4. Nous calculons $(Q_{0,25}(X) - 1,5 \times \text{EIQ}(X))$ et nous cherchons la première observation x_b au delà de la limite $(Q_{0,25}(X) - 1,5 \times \text{EIQ}(X))$, soit $x_b = \min \{x_i | x_i \geq Q_{0,25}(X) - 1,5 \times \text{EIQ}(X)\}$.
5. Nous traçons deux lignes allant des milieux des largeurs du rectangle aux valeurs x_b et x_h .
6. Toute valeur qui n'est pas comprise entre les extrémités des moustaches est représentée par un symbole particulier, une étoile en général. Il s'agit d'une valeur qui est fortement éloignée du reste de la distribution.

3. Caractéristiques d'asymétrie

Nous rappelons que le **moment centré d'ordre r** d'une distribution statistique simple, discrète ou continue, non groupée est égal à :

$${}_{\mu}m_r(X) = \sum_{i=1}^p (x_i - \mu(X))^r f_i.$$

 Fiche 101

Définition 102.1

Le **coefficient d'asymétrie de Fisher** d'une distribution statistique simple, discrète ou continue, *non groupée*, est la quantité $\gamma_1(X)$ définie par :

$$\gamma_1(X) = \frac{{}_{\mu}m_3(X)}{({}_{\mu}m_2(X))^{3/2}} = \frac{{}_{\mu}m_3(X)}{\sigma^3(X)}.$$

► Coefficient d'asymétrie de Pearson

Définition 102.2

Le **coefficient d'asymétrie de Pearson** d'une distribution statistique simple, discrète ou continue, *non groupée*, est la quantité $\beta_1(X)$ définie par :

$$\beta_1(X) = \frac{({}_{\mu}m_3(X))^2}{({}_{\mu}m_2(X))^3} = \frac{({}_{\mu}m_3(X))^2}{\sigma^6(X)}.$$



Il faut noter la relation suivante : $\beta_1(X) = \gamma_1^2(X)$.

Pour une distribution statistique simple *groupée* en classes, il n'est pas nécessaire d'utiliser de correction pour calculer la valeur du coefficient d'asymétrie de Fisher ou de Pearson.

Définition 102.3

Une distribution statistique simple, discrète ou continue, pour laquelle

- $\gamma_1 < 0$ est une distribution ayant une queue de **distribution étalée vers la gauche** ;
- $\gamma_1 = 0$ est une **distribution symétrique** ;
- $\gamma_1 > 0$ est une distribution ayant une queue de **distribution étalée vers la droite**.

4. Caractéristiques d'aplatissement

Définition 102.4

Le **coefficient d'aplatissement de Fisher**, ou excès d'aplatissement, d'une distribution statistique simple, discrète ou continue, *non groupée*, est la quantité $\gamma_2(X)$ définie par :

$$\gamma_2(X) = \frac{\mu m_4(X)}{(\mu m_2(X))^2} - 3 = \frac{\mu m_4(X)}{\sigma^4(X)} - 3.$$

Définition 102.5

Le **coefficient d'aplatissement de Pearson** d'une distribution statistique simple, discrète ou continue, *non groupée*, est la quantité $\beta_2(X)$ définie par :

$$\beta_2(X) = \frac{\mu m_4(X)}{(\mu m_2(X))^2} = \frac{\mu m_4(X)}{\sigma^4(X)} = \gamma_2(X) + 3.$$



Point méthode

Pour une distribution statistique simple *groupée* en classes d'amplitudes toutes égales à a , la valeur du coefficient d'aplatissement de Fisher ou de Pearson se calcule généralement en utilisant la correction de Sheppard pour le terme $\mu m_2(X)$ et une correction similaire à celle-ci pour le terme $\mu m_4(X)$:

$$\begin{aligned}\mu m_4(X) &= \sum_{i=1}^p \left(\frac{a_i + a_{i+1}}{2} - \mu(X) \right)^4 f_i - \frac{a^2}{2} \sum_{i=1}^p \left(\frac{a_i + a_{i+1}}{2} - \mu(X) \right)^2 f_i + \frac{7a^4}{240} \\ &= \frac{1}{N} \sum_{i=1}^p \left(\frac{a_i + a_{i+1}}{2} - \mu(X) \right)^4 n_i - \frac{a^2}{2} \frac{1}{N} \sum_{i=1}^p \left(\frac{a_i + a_{i+1}}{2} - \mu(X) \right)^2 n_i + \frac{7a^4}{240}.\end{aligned}$$

Définition 102.6

Une distribution statistique simple, discrète ou continue, pour laquelle

- $\gamma_2 < 0$ ou $\beta_2 < 3$ est une **distribution platikurtique** ;
- $\gamma_2 = 0$ ou $\beta_2 = 3$ est une **distribution mésokurtique** ;
- $\gamma_2 > 0$ ou $\beta_2 > 3$ est une **distribution leptokurtique**.



Les coefficients d'asymétrie et d'aplatissement sont souvent utilisés pour évaluer la normalité d'une distribution statistique.

Application

1. Le temps de travail des salariés en Europe en 2010

- Établir la boîte à moustaches, en symbolisant la moyenne par un triangle, pour la première application de la Fiche 100.
- Calculer les coefficients d'asymétrie de Fisher et de Pearson.
- Calculer les coefficients d'aplatissement de Fisher et de Pearson.

2. Le temps de travail des non-salariés en Europe en 2010

Établir la boîte à moustaches, en symbolisant la moyenne par un triangle, pour le temps de travail des non-salariés en Europe dont les données sont reproduites dans le Tableau 100.1 de la Fiche 100.

1.a. La Figure 102.1 correspond à la boîte à moustaches demandée.

b. Le coefficient d'asymétrie de Fisher est $-0,43$; celui de Pearson est $0,19$.

c. Le coefficient d'aplatissement de Fisher est $-0,11$; celui de Pearson est $2,89$.

2. La Figure 102.2 correspond à la boîte à moustaches demandée.

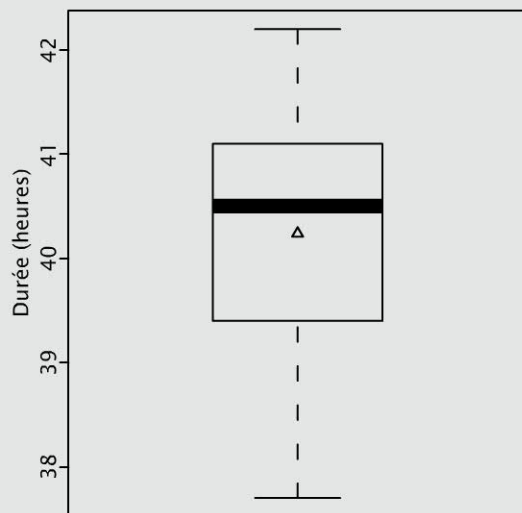


Figure 102.1

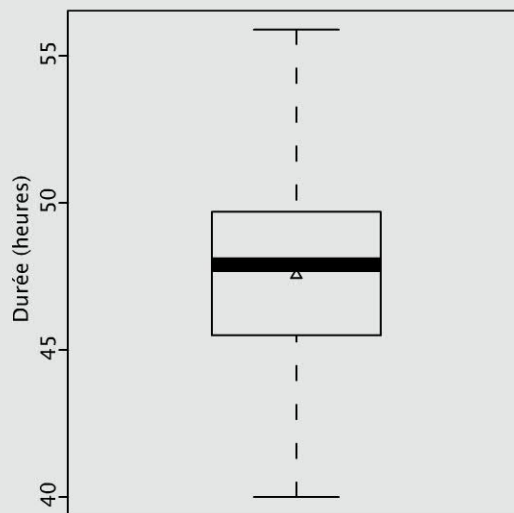


Figure 102.2

Pour s'entraîner (solutions p. 552)

102.1 Pour chacun des trois taux de salariés à temps partiel en Europe du Tableau 100.1 de la Fiche 100, répondre aux questions suivantes.

1. Établir la boîte à moustaches.
2. Calculer les coefficients d'asymétrie de Fisher et de Pearson.
3. Calculer les coefficients d'aplatissement de Fisher et de Pearson.

102.2 Pour la seconde application de cette Fiche, calculer les caractéristiques de formes suivantes :

1. les coefficients d'asymétrie de Fisher et de Pearson.

2. les coefficients d'aplatissement de Fisher et de Pearson.

102.3 Nous avons regroupé en classes, de même amplitude, les caractères qui ont été introduits avec le nouveau jeu de données sur le Travail en Europe à la Fiche 100 et que vous avez déjà étudiés dans cette fiche. Les distributions groupées obtenues ont été reproduites à la Fiche 101. Pour chacune de ces distributions groupées, calculer les caractéristiques de forme suivantes :

1. les coefficients d'asymétrie de Fisher et de Pearson.
2. les coefficients d'aplatissement de Fisher et de Pearson.

1. Présentation des données

Soit (X, Y) un couple de caractères défini sur une population d'effectif N . Notons $x_1 < x_2 < \dots < x_p$ et $y_1 < y_2 < \dots < y_q$ les valeurs distinctes observées pour X et Y et ordonnées dans l'ordre croissant.

2. Distribution conjointe

Définitions 103.1

- L'**effectif du couple** (x_i, y_j) est égal au nombre $n_{i,j}$ de couples de valeurs égaux à (x_i, y_j) , c'est à dire le cardinal de $(X, Y)^{-1}(x_i, y_j)$.
- La **fréquence du couple** (x_i, y_j) est égale au nombre $f_{i,j} = n_{i,j}/N$.
- $((x_i, y_j), n_{i,j})_{i=1,\dots,p, j=1,\dots,q}$ et $((x_i, y_j), f_{i,j})_{i=1,\dots,p, j=1,\dots,q}$ sont la **distribution conjointe des effectifs** et la **distribution conjointe des fréquences** du couple (X, Y) .
- Le **tableau de contingence** est un tableau qui croise les différentes modalités des deux caractères.

$X \backslash Y$	y_1	$\dots$	y_j	$\dots$	y_q	Effectifs $n_{i,\bullet}$
x_1	$n_{1,1}$	$\dots$	$n_{1,j}$	$\dots$	$n_{1,q}$	$n_{1,\bullet}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_p	$n_{p,1}$	$\dots$	$n_{p,j}$	$\dots$	$n_{p,q}$	$n_{p,\bullet}$
Effectifs $n_{\bullet,j}$	$n_{\bullet,1}$	$\dots$	$n_{\bullet,j}$	$\dots$	$n_{\bullet,q}$	$n_{\bullet,\bullet}$

Le nombre entier $n_{\bullet,\bullet}$ est égal à l'effectif de la population N .



Ce tableau représente la distribution conjointe des effectifs du couple (X, Y) .

3. Distributions marginales

Définition 103.2

L'**effectif marginal** de x_i (respectivement de y_j) est égal au nombre $n_{i,\bullet} = \sum_{j=1}^q n_{i,j}$ (respectivement $n_{\bullet,j} = \sum_{i=1}^p n_{i,j}$) et la **fréquence marginale** de x_i (respectivement de y_j) au nombre $f_{i,\bullet} = \sum_{j=1}^q f_{i,j}$ (respectivement $f_{\bullet,j} = \sum_{i=1}^p f_{i,j}$).

$(x_i, n_{i,\bullet})_{i=1,\dots,p}$ et $(y_j, n_{\bullet,j})_{j=1,\dots,q}$ (respectivement $(x_i, f_{i,\bullet})_{i=1,\dots,p}$ et $(y_j, f_{\bullet,j})_{j=1,\dots,q}$) sont appelées les **distributions marginales des effectifs** (respectivement **des fréquences**) de X et de Y .

Les **moyennes marginales** des caractères X et Y sont les quantités respectivement définies par :

$$\mu(X) = \frac{1}{N} \sum_{i=1}^p n_{i,\bullet} x_i = \sum_{i=1}^p f_{i,\bullet} x_i \quad \text{et} \quad \mu(Y) = \frac{1}{N} \sum_{j=1}^q n_{\bullet,j} y_j = \sum_{j=1}^q f_{\bullet,j} y_j.$$

Les **variances marginales** des caractères X et Y sont les quantités respectivement définies par :

$$\sigma^2(X) = \frac{1}{N} \sum_{i=1}^p n_{i,\bullet} (x_i - \mu(X))^2 = \sum_{i=1}^p f_{i,\bullet} x_i^2 - \mu(X)^2$$

et

$$\sigma^2(Y) = \frac{1}{N} \sum_{j=1}^q n_{\bullet,j} (y_j - \mu(Y))^2 = \sum_{j=1}^q f_{\bullet,j} y_j^2 - \mu(Y)^2.$$

4. Indépendance et distributions conditionnelles

Définition 103.3

Y est **indépendant** de X si la distribution des fréquences de Y dans la sous-population $\Omega_{i,\bullet}$ des individus pour lesquels la valeur de X est x_i ne dépend pas de i .

L'effectif de la sous-population $\Omega_{i,\bullet}$ est égal à $n_{i,\bullet}$.

Propriété : La relation d'indépendance est **symétrique**. Si X est indépendant de Y alors Y est indépendant de X . Ainsi nous dirons donc plutôt que **X et Y sont indépendants**.

Définition 103.4

La fraction $\frac{n_{i,j}}{n_{i,\bullet}} = \frac{f_{i,j}}{f_{i,\bullet}}$ représente la **fréquence conditionnelle** de la valeur y_j sachant x_i , notée $f_{j|i}$.

La distribution $\left(y_j, \frac{f_{i,j}}{f_{i,\bullet}}\right)_{j=1,\dots,q}$ est appelée la **distribution conditionnelle des fréquences** de Y sachant que $X = x_i$.



1. Nous définissons d'une manière analogue la **distribution conditionnelle des fréquences** de X sachant que $Y = y_j$.
2. Pour que Y soit indépendant de X il faut que la distribution conditionnelle des fréquences de Y sachant que $X = x_i$ ne dépende pas de i .

Définition 103.5

La **moyenne conditionnelle** et la **variance conditionnelle** du caractère Y sachant que $X = x_i$, sont les quantités respectivement définies par :

$$\mu_i(Y) = \frac{1}{n_{i,\bullet}} \sum_{j=1}^q n_{i,j} y_j = \sum_{j=1}^q f_{j|i} y_j,$$

et

$$\sigma_i^2(Y) = \frac{1}{n_{i,\bullet}} \sum_{j=1}^q n_{i,j}(y_j - \mu_i(Y))^2 = \sum_{j=1}^q f_{j|i} y_j^2 - \mu_i(Y)^2.$$

Le caractère formé des moyennes conditionnelles de Y sachant X est noté $\mu(Y|X)$ et celui formé des variances conditionnelles de Y sachant X est noté $\sigma^2(Y|X)$.



Vous pouvez aussi définir la **moyenne conditionnelle** et la **variance conditionnelle** du caractère X sachant que $Y = y_j$.

Application

1. Déterminer la distribution conjointe des temps partiels des hommes (X) et des femmes (Y) (% des emplois totaux) introduites au Tableau 100.1 Fiche 100 et regroupés suivant les classes introduites à la Fiche 101.
2. Calculer les distributions marginales des effectifs de X et de Y .
3. Calculer les distributions marginales des fréquences de X et de Y .
4. Calculer la distribution conditionnelle des fréquences de Y sachant que $X = x_i$.
5. Calculer la distribution conditionnelle des fréquences de X sachant que $Y = y_i$.

1. La distribution conjointe des temps partiels des hommes (X) et des femmes (Y) est :

		Temps partiels femmes (% emplois totaux)							
		]0; 10]	]10; 20]	]20; 30]	]30; 40]	]40; 50]	]50; 60]	]60; 70]	]70; 80]
Temps partiels hommes (% emplois totaux)	]0; 10]	5	8	3	2	3	0	0	0
	]10; 20]	0	1	0	2	2	0	0	0
	]20; 30]	0	0	0	0	0	0	0	1

2. Nous retrouvons les distributions que nous avons obtenues à la Fiche 101.

Temps partiels hommes (% emplois totaux)							
]0; 10]	]10; 20]	]20; 30]	]30; 40]	]40; 50]	]50; 60]	]60; 70]	]70; 80]
21	5	1	0	0	0	0	0

Temps partiels femmes (% emplois totaux)							
]0; 10]	]10; 20]	]20; 30]	]30; 40]	]40; 50]	]50; 60]	]60; 70]	]70; 80]
5	9	3	4	5	0	0	1

3. Les distributions marginales des fréquences sont les suivantes :

Temps partiels hommes (% emplois totaux)		
]0; 10]	]10; 20]	]20; 30]
$\frac{21}{27}$	$\frac{5}{27}$	$\frac{1}{27}$

Temps partiels femmes (% emplois totaux)							
]0; 10]	]10; 20]	]20; 30]	]30; 40]	]40; 50]	]50; 60]	]60; 70]	]70; 80]
$\frac{5}{27}$	$\frac{9}{27}$	$\frac{3}{27}$	$\frac{4}{27}$	$\frac{5}{27}$	0	0	$\frac{1}{27}$

4. La distribution conditionnelle des fréquences de Y sachant que $X = x_i$ est la suivante :

		Temps partiels femmes (% emplois totaux)							
		]0; 10]	]10; 20]	]20; 30]	]30; 40]	]40; 50]	]50; 60]	]60; 70]	]70; 80]
Temps partiels hommes (%) emplois totaux)	]0;10]	$\frac{5}{21}$	$\frac{8}{21}$	$\frac{3}{21}$	$\frac{2}{21}$	$\frac{3}{21}$	0	0	0
	]10;20]	0	$\frac{1}{5}$	0	$\frac{2}{5}$	$\frac{2}{5}$	0	0	0
	]20;30]	0	0	0	0	0	0	0	1

Cette distribution conditionnelle dépend de x_i : les deux caractères ne sont pas indépendants.

5. La distribution conditionnelle des fréquences de X sachant que $Y = y_i$ est

		Temps partiels femmes (% emplois totaux)							
		[0; 10]	[10; 20]	[20; 30]	[30; 40]	[40; 50]	[50; 60]	[60; 70]	[70; 80]
Temps partiels	[0;10]	1	$\frac{8}{9}$	1	$\frac{1}{2}$	$\frac{3}{5}$	0	0	0
hommes (%)	[10;20]	0	$\frac{1}{9}$	0	$\frac{1}{2}$	$\frac{2}{5}$	0	0	0
emplois totaux)	[20;30]	0	0	0	0	0	0	0	1



Nous constatons, comme cela était prévisible puisque la relation d'indépendance est symétrique, que cette distribution conditionnelle des fréquences dépend de y_i , ce qui confirme que les deux caractères ne sont pas indépendants.

Pour s'entraîner (solutions p. 553)

103.1

1. Montrer que la distribution conjointe des temps de travail des salariés (X) et des non-salariés (Y) (en heures par semaines) regroupés suivant les classes introduites à la Fiche 101 est celle dans le tableau ci-dessous.

2. Calculer les distributions marginales des effectifs et des fréquences de X et de Y .

3. Déterminer la distribution conditionnelle des fréquences de X sachant que $Y = y_i$. Les caractères X et Y sont-elles indépendants ?

4. Déterminer la distribution conditionnelle des fréquences de Y sachant que $X = x_i$.

		Temps de travail non-salariés (heures/semaines)				
Temps de travail		[35; 40]	[40; 45]	[45; 50]	[50; 55]	[55; 60]
salariés	[35; 40]	0	1	5	3	1
(heures/semaines)	[40; 45]	1	3	11	2	0

Distribution conjointe des temps de travail des salariés et des non-salariés (en heures par semaines).

1. Présentation

Le mode de représentation graphique d'une série statistique bivariée dépend de la nature de chacun des caractères : qualitatifs, quantitatifs discrets ou continus. Il existe une grande diversité de représentations graphiques de ces distributions statistiques. Nous exposerons ci-dessous les plus utilisées.

2. Cas de deux caractères qualitatifs

► Distributions marginales

Lorsque nous croisons deux caractères qualitatifs, la représentation graphique d'une distribution marginale est réalisée à l'aide d'un diagramme à barres verticales, voir Figure 104.1, (ou horizontales) ou d'un diagramme circulaire.

► Distributions conditionnelles

La représentation graphique des distributions conditionnelles est réalisée à l'aide d'un diagramme à barres empilées, voir Figure 104.2, ou d'un diagramme à barres juxtaposées, voir Figure 104.3. Ce type de représentation a pour avantage de permettre de comparer rapidement les différentes distributions, en particulier savoir si elles sont homogènes ou non.

3. Cas d'un caractère quantitatif et d'un caractère qualitatif

► Distributions marginales

Pour le caractère quantitatif, discret ou continu, la représentation graphique de la distribution marginale est réalisée à l'aide d'un diagramme en bâtons ou d'un histogramme, voir Figure 105.1, selon que la distribution est non groupée ou groupée.

Pour le caractère qualitatif, la représentation graphique de la distribution marginale est réalisée à l'aide d'un diagramme à barres verticales (ou horizontales) ou d'un diagramme circulaire, voir Figure 105.2.

► Distributions conditionnelles

Sachant les modalités du caractère qualitatif, la représentation graphique des distributions conditionnelles est réalisée à l'aide de diagrammes en bâtons ou d'histogrammes, voir Figure 105.3, selon que la distribution quantitative, discrète ou continue, est non groupée ou groupée (nous pouvons représenter autant de diagrammes que le caractère qualitatif possède de modalités).

Sachant les classes du caractère quantitatif, la représentation graphique des distributions conditionnelles est réalisée à l'aide d'un diagramme à barres juxtaposées, voir Figure 105.4, ou d'un diagramme à barres empilées (nous pouvons représenter autant de diagrammes que le caractère quantitatif possède de classes).

4. Cas de deux caractères quantitatifs

► Distributions marginales

La représentation graphique est réalisée à l'aide de diagrammes en bâtons ou d'histogrammes, voir Figure 104.4, selon que les caractères sont quantitatifs discrets ou continus.

► Distributions conditionnelles

La représentation graphique est réalisée à l'aide de diagrammes en bâtons ou d'histogrammes, voir Figure 104.5 ou Figure 104.6, selon que les caractères sont quantitatifs discrets ou continus.

► Distribution conjointe

Lorsque nous disposons d'une distribution groupée et de ce fait du tableau de contingence associé, la représentation graphique de la distribution jointe est réalisée à l'aide d'un stéréogramme (représentation en 3D), voir Figure 104.8. Dans le cas de données brutes, nous représentons le diagramme de dispersion : dans un repère cartésien, chaque observation (x_k, y_k) est représentée par un point M_k de coordonnées (x_k, y_k) , voir Figure 104.7. Ce type de graphique est particulièrement important pour l'étude d'une éventuelle liaison entre les deux caractères.



Lorsque nous disposons d'un tableau de contingence croisant deux caractères quantitatifs, il est possible de représenter le diagramme de dispersion. Pour cela, nous remplaçons les données brutes par les centres de classes, puis nous représentons pour, chaque point, un cercle coloré centré sur le point dont le diamètre est proportionnel à la fréquence $f_{i,j}$.

Application

1. Considérons la répartition des étudiants français ayant bénéficié du programme Erasmus de 2002/2003 à 2007/2008 par rapport aux pays d'accueil. Celle-ci a été introduite dans l'application de la Fiche 96.

- a. Établir une représentation graphique des effectifs marginaux.
- b. Établir une représentation graphique des fréquences conditionnelles.

2. Nous nous intéressons à la distribution conjointe des temps partiels des hommes (X) et des femmes (Y) (% des emplois totaux). Celle-ci a été introduite dans l'application de la Fiche 103.

- a. Représenter les distributions marginales de X et Y à l'aide des graphiques adaptés.
- b. Représenter les distributions conditionnelles de X et Y à l'aide des graphiques adaptés.
- c. Représenter conjointement les données brutes de X et Y à l'aide du graphique adapté.
- d. Représenter la distribution conjointe de X et Y , obtenue à la Fiche 103 à partir des classes introduites à la Fiche 101, à l'aide du graphique adapté.

1.a. Les représentations graphiques sont les diagrammes en barres verticales reproduits à la Figure 104.1.

b. Les représentations graphiques sont les diagrammes en barres verticales juxtaposées reproduits à la Figure 104.3 ou en barres empilées reproduits à la Figure 104.2.

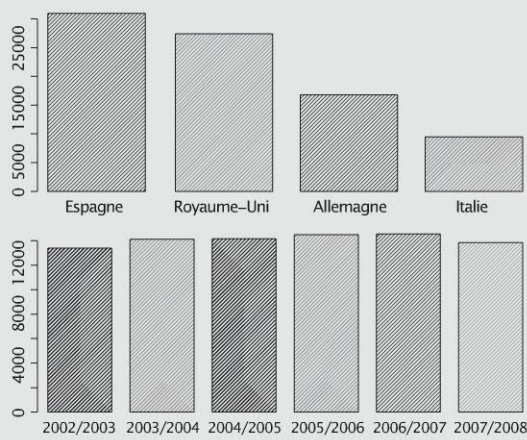


Figure 104.1

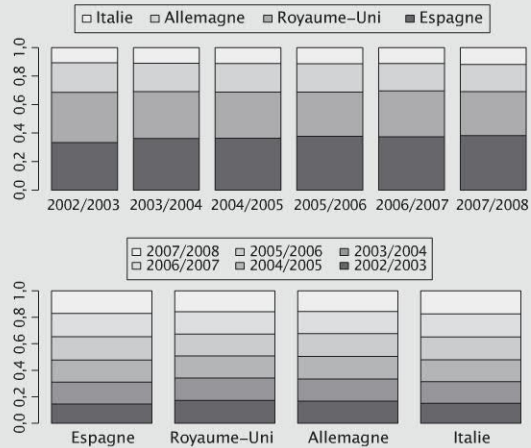


Figure 104.2

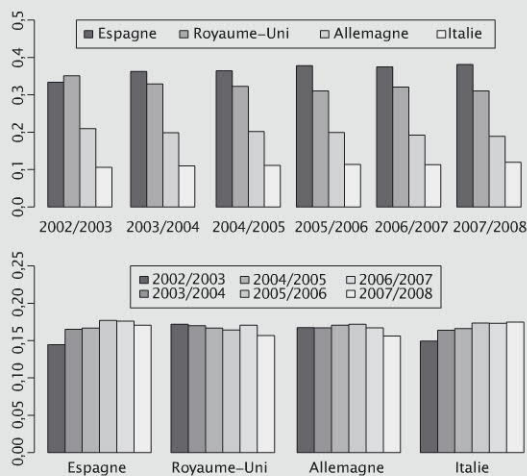


Figure 104.3

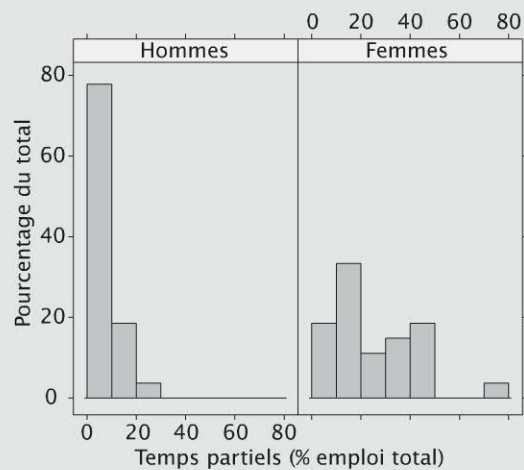


Figure 104.4

2.a. Les représentations graphiques sont les histogrammes reproduits à la Figure 104.4.

b. Les représentations graphiques sont les histogrammes reproduits aux Figures 104.2 et 104.3.

c. La représentation graphique est un diagramme de dispersion reproduit à la Figure 104.7.

d. La représentation graphique est un stéréogramme reproduit à la Figure 104.8.

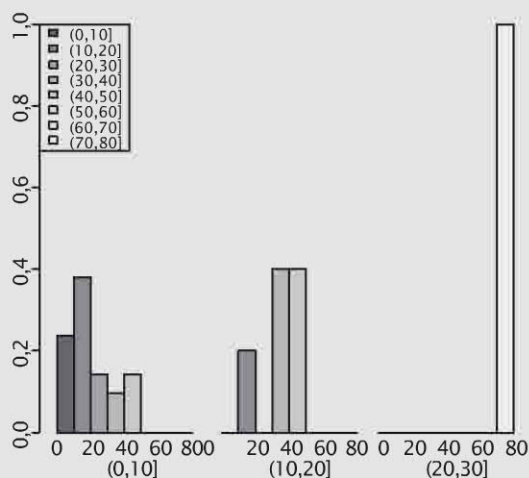


Figure 104.5

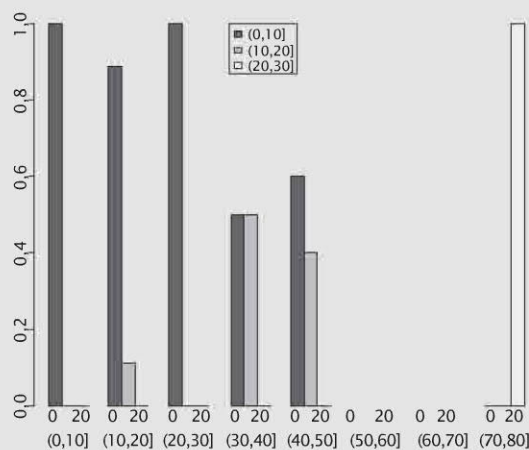


Figure 104.6

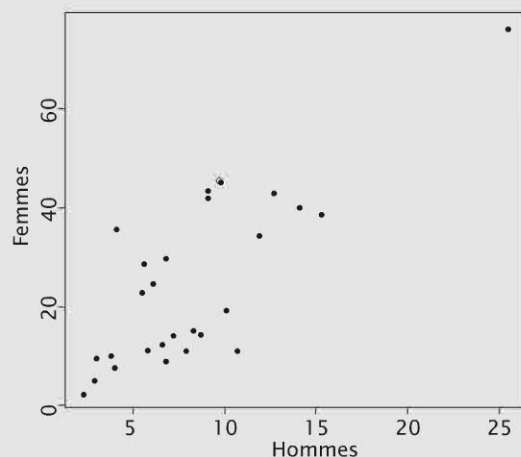


Figure 104.7

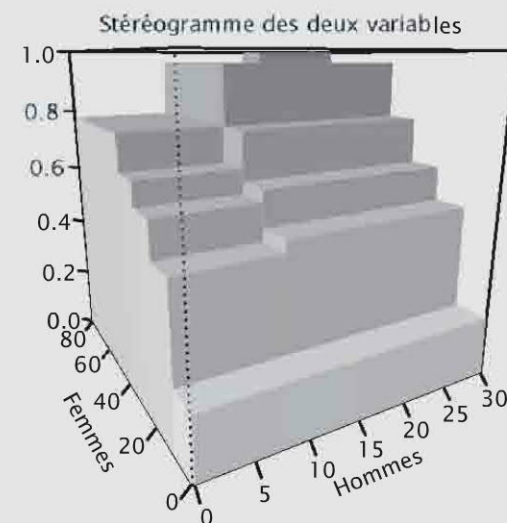


Figure 104.8

Pour s'entraîner

104.1



Nous nous intéressons à la distribution conjointe du nombre de sièges au parlement et de voix au conseil pour les 27 pays de l'Union Européenne (Tableau 99.1).

1. Représenter les distributions marginales de X et Y à l'aide des graphiques adaptés.
2. Représenter les distributions conditionnelles de X et Y à l'aide des graphiques adaptés.

104.2



Nous nous intéressons à la distribution conjointe de la durée de travail des salariés (X) et des non-salariés (Y) (en heures) pour les 27 pays de l'Union Européenne introduite à la Fiche 100.

1. Représenter les distributions marginales de X et Y à l'aide des graphiques adaptés.
2. Représenter les distributions conditionnelles de X et Y à l'aide des graphiques adaptés.
3. Représenter conjointement les données brutes de X et Y à l'aide du graphique adapté.

Mesures de liaison entre deux caractères

1. Liaison entre deux caractères quantitatifs

Définition 105.1

La **covariance** de la série statistique double du couple de caractères quantitatifs (X, Y) est le nombre défini par :

$$\mathcal{C}ov(X, Y) = \sum_{i,j} (x_i - \mu(X))(y_j - \mu(Y))f_{i,j}.$$

Proposition 105.1

En développant l'égalité définissant la covariance, nous obtenons la relation suivante :

$$\mathcal{C}ov(X, Y) = \mu(XY) - \mu(X)\mu(Y)$$

avec $\mu(XY)$, la moyenne de XY . Elle est égale à :

$$\mu(XY) = \sum_{i,j} x_i y_j f_{i,j},$$

où la fréquence du couple, $f_{i,j}$, a été définie à la Fiche 103, la moyenne du caractère X , $\mu(X)$, et la moyenne du caractère Y , $\mu(Y)$, ont été définies à la Fiche 100.

 Fiches 100
et 103



Si les deux caractères X et Y sont indépendants, alors la covariance du couple (X, Y) est nulle. **La réciproque de cette proposition est fausse.** Il existe des exemples pour lesquels $\mathcal{C}ov(X, Y) = 0$ sans que les caractères soient indépendants. Par conséquent, la covariance n'est pas une mesure complète de la dépendance. En fait, la covariance mesure uniquement la dépendance linéaire. L'intensité de cette dépendance est donnée par un nombre sans dimension : le coefficient de corrélation linéaire.

 Fiche 103

Définition 105.2

Le **coefficient de corrélation linéaire** de la série statistique double du couple de caractères quantitatifs (X, Y) est le quotient défini par :

$$\mathcal{R}(X, Y) = \frac{\mathcal{C}ov(X, Y)}{\sigma(X)\sigma(Y)}.$$



1. Si les deux caractères X et Y sont indépendants, alors le coefficient de corrélation linéaire est nul. **La réciproque de cette proposition est fausse.**
2. Nous avons toujours $|\mathcal{R}(X, Y)| \leq 1$.
3. Si $|\mathcal{R}(X, Y)| = 1$, il existe une relation linéaire exacte entre les caractères X et Y .

2. Liaison entre deux caractères qualitatifs

Définition 105.3

Le **coefficient d'association** mesure le degré de liaison entre deux caractères qualitatifs X et Y . Il est noté $\chi^2(X, Y)$ et est égal à :

$$\chi^2(X, Y) = \sum_{i,j} \frac{\left(n_{i,j} - \frac{n_{i,\bullet} \times n_{\bullet,j}}{n_{\bullet,\bullet}} \right)^2}{\frac{n_{i,\bullet} \times n_{\bullet,j}}{n_{\bullet,\bullet}}}.$$



Le coefficient d'association $\chi^2(X, Y)$ est nul lorsque les effectifs réels $n_{i,j}$ et les effectifs calculés sous hypothèse d'indépendance $n_{i,\bullet} n_{\bullet,j} / n_{\bullet,\bullet}$ des couples (x_i, y_j) sont égaux.

Définition 105.4

Le **coefficient d'association de Pearson**, noté $\phi^2(X, Y)$, est égal à : $\phi^2(X, Y) = \chi^2(X, Y) / n_{\bullet,\bullet}$.



Ce coefficient $\phi^2(X, Y)$ ne dépend plus de l'effectif de la population qu'il sert à étudier et est nul lorsque le coefficient d'association $\chi^2(X, Y)$ est nul.

3. Liaison entre un caractère qualitatif et un caractère quantitatif

Soient Y un caractère quantitatif, X un caractère qualitatif, $n_{i,\bullet}$ l'effectif marginal et $f_{i,\bullet}$ la fréquence marginale de x_i , $\mu(Y)$ la moyenne de Y , $\mu_i(Y)$ la moyenne conditionnelle et $\sigma_i^2(Y)$ la variance conditionnelle de Y sachant que $X = x_i$. Enfin, $\mu(Y|X)$ est la distribution des moyennes conditionnelles de Y sachant X et $\sigma^2(Y|X)$ celle des variances conditionnelles.

Fiche 103

Définition 105.5

La **variance inter-modalités** ou **variance expliquée par le caractère qualitatif** est égale à :

$$\sigma^2(\mu(Y|X)) = 1/n_{\bullet,\bullet} \sum_{i=1}^p n_{i,\bullet} (\mu_i(Y) - \mu(Y))^2 = \sum_{i=1}^p f_{i,\bullet} (\mu_i(Y) - \mu(Y))^2.$$

Définition 105.6

La **variance intra-modalités** ou **variance résiduelle** est égale à :

$$\mu(\sigma^2(Y|X)) = 1/n_{\bullet,\bullet} \sum_{i=1}^p n_{i,\bullet} \sigma_i^2(Y) = \sum_{i=1}^p f_{i,\bullet} \sigma_i^2(Y).$$

Théorème 105.1 Décomposition de la variance

$$\sigma^2(Y) = \mu(\sigma^2(Y|X)) + \sigma^2(\mu(Y|X)).$$

Définition 105.7

La mesure, ici non symétrique, de la liaison entre le caractère qualitatif X et le caractère quantitatif Y est le **rapport de corrélation**, noté $\eta_{Y|X}^2$, défini par :

$$\eta_{Y|X}^2 = \sigma^2(\mu(Y|X)) / \sigma^2(Y).$$



1. Le rapport de corrélation est toujours compris entre 0 et 1.
2. Si les deux caractères X et Y sont indépendants, alors le rapport de corrélation $\eta_{Y|X}^2$ est nul.
La réciproque de cette proposition est fausse.
3. Il ne faut pas confondre ce coefficient avec le coefficient de corrélation linéaire. Il existe néanmoins une relation entre ces deux quantités : $\mathcal{R}^2(X, Y) \leq \eta_{Y|X}^2$.

Application

1. Nous nous intéressons à la distribution conjointe de la durée de travail des salariés et des non-salariés (en heures) pour les 27 pays de l'Union Européenne introduite à la Fiche 100.

- a. Calculer la covariance et le coefficient de corrélation linéaire entre ces deux caractères.
- b. Les caractères sont-ils liés ? Commenter.

2. Considérons la répartition des étudiants français ayant bénéficié du programme Erasmus de 2002/2003 à 2007/2008 par rapport aux pays d'accueil. Celle-ci a été introduite dans l'application de la Fiche 96.

- a. Donner les valeurs des effectifs sous l'hypothèse d'indépendance entre les deux caractères.
- b. Calculer la valeur du coefficient d'association et du coefficient d'association de Pearson.
- c. Les caractères sont-ils liés ? Commenter.

3. Considérons le taux de temps partiels (Y) en fonction de la variable type (X) qui peut prendre trois modalités : ensemble, hommes et femmes, introduit à la Fiche 100.

- a. Représenter les distributions marginales et conditionnelles des deux caractères.
- b. Calculer le rapport de corrélation entre ces deux caractères.
- c. Les caractères sont-ils liés ?
- d. Calculer la variance inter-modalités et la variance intra-modalités.

1.a. $\text{Cov}(X, Y) = -0,554 \text{ h}^2$, $\sigma(X) = 1,036 \text{ h}$, $\sigma(Y) = 3,671 \text{ h}$ et $\mathcal{R}(X, Y) = -0,146$.

b. $\mathcal{R}(X, Y) < 0$ donc les deux caractères sont liés et évoluent en sens contraire.

2.a. Les effectifs calculés sous hypothèse d'indépendance des couples (x_i, y_j) sont $n_{i\bullet}n_{\bullet j}/n_{\bullet\bullet}$, soit :

	2002/2003	2003/2004	2004/2005	2005/2006	2006/2007	2007/2008	$n_{i\bullet}$
Espagne	4904,0	5168,3	5184,8	5309,6	5331,2	5070,2	30968,0
Royaume-Uni	4337,7	4571,5	4586,1	4696,5	4715,6	4484,7	27392,0
Allemagne	2656,8	2799,9	2808,9	2876,5	2888,2	2746,8	16777,0
Italie	1499,5	1580,3	1585,3	1623,5	1630,1	1550,3	9469,0
$n_{\bullet j}$	13398,0	14120,0	14165,0	14506,0	14565,0	13852,0	84606,0



Fiche 113

b. Le coefficient d'association est $\chi^2(X, Y) = 136,3$, celui de Pearson est $\phi^2(X, Y) = 0,00161$.

c. Les caractères sont liés car les deux valeurs précédentes sont différentes de 0 : les préférences des étudiants ne sont pas restées identiques de 2002/2003 à 2007/2008.

3.a. Les Figures 105.1, 105.2, 105.3, 105.4 correspondent aux représentations graphiques demandées.

b. Le rapport de corrélation est $\eta_{Y|X}^2 = 0,252$.

c. Le rapport de corrélation est non nul, les caractères sont liés.

d. Nous avons $\sigma^2(\mu(Y|X)) = 45,35$ et $\mu(\sigma^2(Y|X)) = 134,89$.

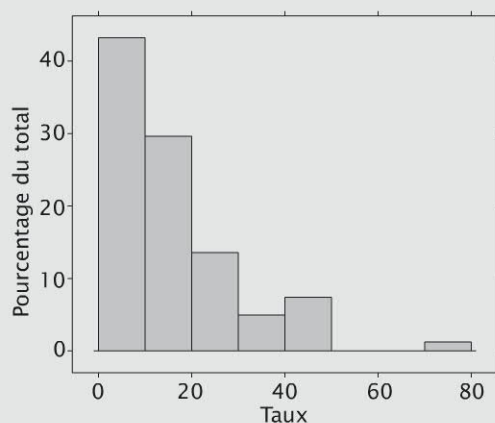


Figure 105.1

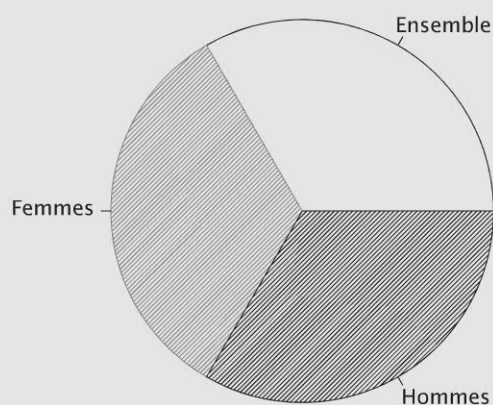


Figure 105.2

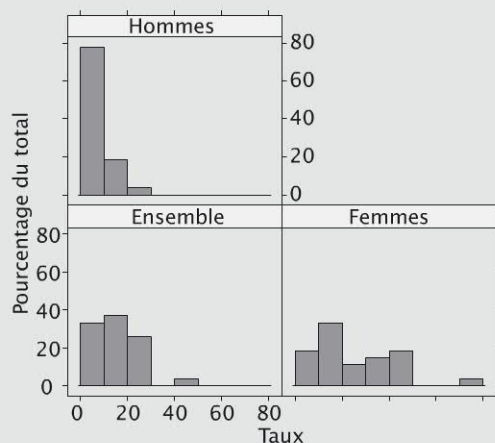


Figure 105.3

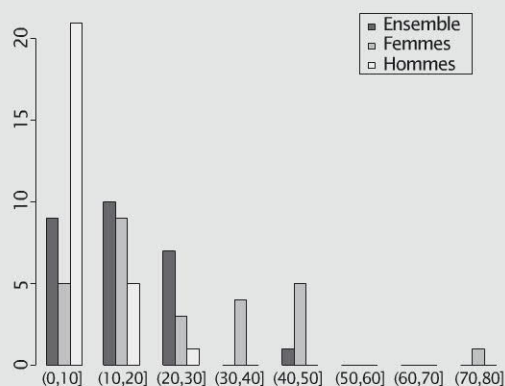


Figure 105.4

Pour s'entraîner (solutions p. 553)

105.1 Nous nous intéressons à la distribution conjointe du nombre de sièges au parlement et de voix au conseil pour les 27 pays de l'Union Européenne (Tableau 99.1 Fiche 99).

1. Calculer la covariance et le coefficient de corrélation linéaire entre ces deux caractères.
2. Les caractères sont-ils liés ? Commenter.

1. Introduction

Le **problème de l'estimation** est l'impossibilité de connaître exactement la valeur d'un **paramètre inconnu** noté θ . Ce problème est très général et a des aspects distincts. Les observations, par exemple obtenues à partir d'une méthode d'échantillonnage, permettent de construire **une estimation de θ** . Ainsi chaque observation est la valeur d'une variable aléatoire X dont la loi dépend de θ . Cela revient à représenter le phénomène inconnu étudié à l'aide un **modèle probabiliste**. Ce dernier est complété par un **modèle d'échantillonnage** décrivant la manière dont les observations sont recueillies.

2. Échantillonnage

Définition 106.1

Soit X une variable aléatoire réelle. Un **échantillon aléatoire** d'effectif $n \geq 1$ est un vecteur aléatoire $X_n = (X_1, X_2, \dots, X_n)$ à n composantes qui sont n variables aléatoires indépendantes suivant la même loi que X , appelée **variable aléatoire parente**.

 Fiches 75
et 86



Pour des raisons de commodité, nous avons supposé que les X_i sont mutuellement indépendantes. Dans certains cas, l'indépendance deux à deux sera suffisante.

 Fiche 90

Définition 106.2

Toute variable aléatoire T , fonction de l'échantillon aléatoire $X_n = (X_1, X_2, \dots, X_n)$, est appelée **statistique de l'échantillon**.



1. Une statistique peut être à valeurs dans $\mathbb{R}$ ou $\mathbb{R}^p$. Dans le dernier cas, nous parlerons de **statistique vectorielle**.
2. La difficulté de cette notion est la suivante : nous avons une double conception, qui est la base de la statistique mathématique. Les valeurs observées $(x_1, x_2, \dots, x_n)$ (notons que ce sont des minuscules) constituent n réalisations indépendantes d'une variable aléatoire X ou encore, une réalisation unique du vecteur aléatoire $X_n = (X_1, X_2, \dots, X_n)$ à n composantes où les X_i sont n variables aléatoires indépendantes et de même loi.

Définition 106.3

La **théorie de l'échantillonnage** se propose d'étudier les propriétés du vecteur aléatoire à n composantes et des caractéristiques le résumant, encore appelées statistiques, à partir de la distribution supposée connue de la variable parente X , et d'étudier en particulier ce qui se passe lorsque la taille de l'échantillon est de plus en plus élevée.



Une taille d'échantillon relativement élevée est ce qui préoccupe généralement les statisticiens bien que, depuis quelques années, des théories concernant les petits échantillons se développent également.

3. Modèles

Un phénomène aléatoire repéré par une caractéristique vectorielle $\mathbf{X}$ est parfaitement défini par la donnée d'un couple $(D_{\mathbf{X}}, \mathbb{P}_{\theta})$, où $D_{\mathbf{X}}$ est l'ensemble des résultats possibles pour la caractéristique $\mathbf{X}$ et $\mathbb{P}_{\theta}$ la loi de probabilité de $\mathbf{X}$.

 Fiche 75

Définition 106.4

Un **modèle statistique** est défini par la donnée d'une caractéristique vectorielle $\mathbf{X}$ et d'une famille de lois de probabilité de $\mathbf{X}$ notée $(\mathbb{P}_{\theta})_{\theta \in \Theta}$. Il est généralement noté $(D_{\mathbf{X}}, \mathbb{P}_{\theta})_{\theta \in \Theta}$.

Définition 106.5

Un **modèle d'échantillonnage** décrit la manière dont les observations de l'échantillon sont recueillies.

Certains auteurs utilisent parfois la Définition 106.6 pour définir un modèle statistique.

Définition 106.6

Un **modèle probabiliste** complété par un **modèle d'échantillonnage** est souvent appelé un **modèle statistique**.

Souvent le phénomène est connu de façon imparfaite. L'observateur ou l'expérimentateur connaît $D_{\mathbf{X}}$ et une famille $\mathbb{P}$ à laquelle appartient la vraie loi $\mathbb{P}_{\theta}$ régissant le phénomène.

Définition 106.7

Lorsque la famille de lois de probabilité $(\mathbb{P}_{\theta})_{\theta \in \Theta}$ peut être indexée par un paramètre δ dont l'ensemble des valeurs possibles, noté Δ , est une partie de $\mathbb{R}^k$, le modèle est appelé **modèle paramétrique**. Dans le cas contraire, le modèle est appelé **modèle non paramétrique**.

Définition 106.8

Un modèle non paramétrique peut posséder à la fois une composante paramétrique et une composante non paramétrique. Il est alors appelé **modèle semi-paramétrique**.

1. Estimateur et estimation

Définition 107.1

Si $\mathbf{X}_n = (X_1, \dots, X_n)$ est un échantillon aléatoire d'effectif n de loi parente la loi $\mathbb{P}_\theta$ alors nous appelons **estimateur** de θ toute fonction h_n de l'échantillon aléatoire $\mathbf{X}_n$, noté $\widehat{\theta}_n$:

$$\widehat{\theta}_n = h_n(\mathbf{X}_n) = h_n(X_1, \dots, X_n).$$



1. À priori l'estimateur $\widehat{\theta}_n$ est à valeurs dans un ensemble $\widetilde{\Theta}$, contenant au moins Θ , l'ensemble des valeurs possibles de θ .
2. $\widehat{\theta}_n$ est une **variable aléatoire** de loi de probabilité qui dépend de θ .
3. $\widehat{\theta}_n$ peut être univarié ou multivarié.

Définition 107.2

Une fois l'échantillon prélevé, nous disposons de n valeurs observées $x_1, \dots, x_n$, donc d'une valeur $h_n(x_1, \dots, x_n)$ qui est une réalisation de $\widehat{\theta}_n$ et que nous appelons **estimation**.



1. Nous distinguons la variable aléatoire $\widehat{\theta}_n$ de sa valeur observée, notée $\widehat{\theta}_n(x_1, \dots, x_n)$.
2. Nous utiliserons les notations suivantes :
 - (a) $(X_1, \dots, X_n)$ désigne l'échantillon aléatoire de taille n et les n observations ne sont pas encore à disposition.
 - (b) $(x_1, \dots, x_n)$ désigne une réalisation de l'échantillon aléatoire et les n observations sont à disposition.
3. Il faut systématiquement se demander : « suis-je en train de manipuler une variable aléatoire ou l'une de ses réalisations ? »
4. $\mathbb{E}_\theta(\widehat{\theta}_n)$ et $\text{Var}_\theta(\widehat{\theta}_n)$ sont l'espérance et la variance de l'estimateur $\widehat{\theta}_n$ sachant que la vraie valeur du paramètre est θ .

2. Propriétés d'un estimateur

Le **choix d'un estimateur** va reposer sur ses **qualités**. Commençons par rappeler la notion de convergence en probabilité.

Fiche 92

Définition 107.3

Soit $(X_n)_{n \geq 1}$ une suite de variables aléatoires et X une variable aléatoire. La suite $(X_n)_{n \geq 1}$ **converge en probabilité** vers X si, pour tout $a > 0$, $\lim_{n \rightarrow +\infty} \mathbb{P}(|X_n - X| \geq a) = 0$.

Définition 107.4

Un estimateur $\widehat{\theta}_n$ est un **estimateur convergent** s'il converge en probabilité vers θ quand n tend vers l'infini.

Définition 107.5

Le **biais** de $\widehat{\theta}_n$ se définit par $B_\theta(\widehat{\theta}_n) = \mathbb{E}_\theta(\widehat{\theta}_n) - \theta$.



Lorsque le biais est positif, l'estimateur sur-estime la valeur du paramètre ; lorsqu'il est négatif, l'estimateur sous-estime la valeur du paramètre.

Définition 107.6

$\widehat{\theta}_n$ est un **estimateur sans biais** (ou non biaisé) du paramètre θ si $B_\theta(\widehat{\theta}_n) = 0$, c'est-à-dire si $\mathbb{E}_\theta(\widehat{\theta}_n) = \theta$, pour tout $\theta \in \Theta$.

Définition 107.7

Un estimateur $\widehat{\theta}_n$ est **asymptotiquement sans biais** pour θ si $\lim_{n \rightarrow +\infty} \mathbb{E}_\theta(\widehat{\theta}_n) = \theta$.

Définition 107.8

Si $\widehat{\theta}_n$ est un estimateur de θ , nous mesurons la précision de $\widehat{\theta}_n$ par l'**écart quadratique moyen**, noté EQM_θ :

$$EQM_\theta(\widehat{\theta}_n) = \mathbb{E}_\theta((\widehat{\theta}_n - \theta)^2) = \text{Var}_\theta(\widehat{\theta}_n) + B_\theta(\widehat{\theta}_n)^2.$$



Si $\widehat{\theta}_n$ est un estimateur sans biais, c'est-à-dire si $B_\theta(\widehat{\theta}_n) = 0$, alors $EQM_\theta(\widehat{\theta}_n) = \text{Var}_\theta(\widehat{\theta}_n)$.

Corollaire 107.1

Entre deux estimateurs de θ , nous choisissons celui dont l'écart quadratique moyen ou le risque est le plus faible.

Définition 107.9

Un estimateur $\widehat{\theta}_{n,1}$ est **relativement plus efficace** qu'un estimateur $\widehat{\theta}_{n,2}$ s'il est plus précis que le second, c'est-à-dire si pour tout $\theta \in \Theta$:

$$EQM_\theta(\widehat{\theta}_{n,1}) \leq EQM_\theta(\widehat{\theta}_{n,2}).$$

Définition 107.10

Nous appelons **estimateur sans biais optimal** parmi les estimateurs sans biais, un estimateur $\widehat{\theta}_n$ préférable à tout autre au sens de la variance c'est-à-dire l'estimateur le plus efficace parmi tous les estimateurs sans biais car il concentre le plus possible ses valeurs autour du paramètre.

Théorème 107.1

Si un estimateur est asymptotiquement sans biais et que sa variance tend vers zéro quand n tend vers l'infini, alors cet estimateur est convergent.

Application

Soit $(X_1, \dots, X_n)$ un échantillon aléatoire de loi parente la loi $\mathbb{P}_\theta$. Supposons que, pour tout $\theta \in \Theta$, la loi $\mathbb{P}_\theta$ admette une espérance $\mu(\theta)$, notée simplement μ .

1. Montrer que l'estimateur de l'espérance défini par $\widehat{\mu}_n = \frac{1}{n} \sum_{i=1}^n X_i$ est un estimateur sans biais de la moyenne μ . Supposons de plus que, pour tout $\theta \in \Theta$, la loi $\mathbb{P}_\theta$ admette une variance égale à $\sigma^2(\theta) \in \mathbb{R}_+^*$, notée simplement σ^2 .
2. Calculer la variance de l'estimateur $\widehat{\mu}_n$ et montrer qu'elle tend vers 0 quand n tend vers l'infini.
3. Que conclure ?

1. Calculons l'espérance de $\widehat{\mu}_n$ afin de montrer que l'estimateur est sans biais.

$$\mathbb{E}_\theta(\widehat{\mu}_n) = \mathbb{E}_\theta\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n} \sum_{i=1}^n \mathbb{E}_\theta(X_i) = \frac{1}{n} \sum_{i=1}^n \mu = \mu,$$

puisque pour tout entier i de $\{1, \dots, n\}$, $\mathbb{E}_\theta(X_i) = \mu$. Donc l'estimateur $\widehat{\mu}_n$ est sans biais.

2. Calculons la variance de $\widehat{\mu}_n$.

$$\mathbb{V}\text{ar}_\theta(\widehat{\mu}_n) = \mathbb{V}\text{ar}_\theta\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n^2} \mathbb{V}\text{ar}_\theta\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \mathbb{V}\text{ar}_\theta(X_i)$$

car les variables X_i sont supposées indépendantes donc décorrélées et que pour tout entier i de $\{1, \dots, n\}$, $\mathbb{V}\text{ar}_\theta(X_i) = \sigma^2$. Par conséquent, nous avons :

$$\mathbb{V}\text{ar}_\theta(\widehat{\mu}_n) = \frac{1}{n^2} \sum_{i=1}^n \sigma^2 = \frac{\sigma^2}{n}.$$

Donc quand n tend vers l'infini, $\mathbb{V}\text{ar}_\theta(\widehat{\mu}_n)$ tend vers 0.

3. D'après le Théorème 107.1, l'estimateur $\widehat{\mu}_n$ est un estimateur convergent de $\mu = \mathbb{E}_\theta(X_1)$.

Pour s'entraîner

107.1
[www](#)

Soit $(X_1, \dots, X_n)$ un échantillon aléatoire de loi parente la loi $\mathbb{P}_\theta$. Supposons que, pour tout $\theta \in \Theta$, la loi $\mathbb{P}_\theta$ admette une variance $\sigma^2(\theta)$, notée simplement σ^2 .

1.a. Montrer que l'estimateur non corrigé de la variance défini par $S_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \widehat{\mu}_n)^2$ est un estimateur biaisé de la variance σ^2 si

$\sigma^2(\Theta) \neq \{0\}$, c'est-à-dire si les variances des lois $\mathbb{P}_\theta$ ne sont pas nulles pour tout $\theta \in \Theta$.

b. Montrer que S_n^2 est asymptotiquement sans biais.

2. Montrer que l'estimateur corrigé de la variance défini par $S_{n,c}^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \widehat{\mu}_n)^2$ est un estimateur sans biais de la variance σ^2 .

 Fiche 110

1. Estimateur de la méthode des moments

Cette méthode peut être qualifiée de « quasi naturelle » dans la mesure où elle est relativement intuitive. Supposons que nous devons estimer I paramètres notés $\theta_1, \theta_2, \dots, \theta_I$.

Définition 108.1

La **méthode des moments** consiste à choisir des estimateurs $\widehat{\theta}_1, \dots, \widehat{\theta}_I$ issus de la résolution d'un système de I équations en les I paramètres inconnus, obtenu en égalant les I premiers moments (par rapport à l'origine) notés $\mu_i = \mathbb{E}_{(\theta_1, \dots, \theta_I)}(X^i)$, pour i variant de 1 à I et les moments correspondants de l'échantillon $m_i = \frac{1}{n} \sum_{j=1}^n X_j^i$ pour i variant de 1 à I .

Exemple

Dans le cas où un seul paramètre est inconnu, nous devons résoudre l'équation suivante :

$$\mu_1 = \mu = \mathbb{E}_{\theta_1}(X) = m_1 = \frac{1}{n} \sum_{j=1}^n X_j^1 = \bar{X}_n$$

où μ est la moyenne de la population. Si μ est le paramètre inconnu, la solution de l'équation ci-dessus est triviale :

$$\widehat{\mu} = m_1 = \bar{X}_n.$$

2. Estimateur du maximum de vraisemblance

Définition 108.2

La **vraisemblance des observations** $\mathbf{x} = (x_1, \dots, x_n)$ d'un échantillon aléatoire de loi parente la loi de X est définie de la façon suivante :

- si X est une variable aléatoire continue :

$$\theta \in \Theta \mapsto L(x_1, \dots, x_n | \theta) = \prod_{i=1}^n f_X(x_i, \theta),$$

où Θ est l'ensemble des valeurs possibles du paramètre θ et f_X la densité de probabilité de X ;

- si X est une variable aléatoire discrète :

$$\theta \in \Theta \mapsto L(x_1, \dots, x_n | \theta) = \prod_{i=1}^n f_X(x_i, \theta),$$

où Θ est l'ensemble des valeurs possibles du paramètre θ et f_X est la fonction de probabilité de X , $f_X(x_i, \theta) = \mathbb{P}_\theta(X_i = x_i)$.

 Fiche 75

 Fiche 75



Les expressions des vraisemblances ci-dessus ne sont valables que parce que les variables aléatoires $X_1, \dots, X_n$ sont indépendantes par définition d'un **échantillon aléatoire**.

Définition 108.3

Un **estimateur du maximum de vraisemblance** (EMV) du paramètre θ est une statistique de l'échantillon :

$$\begin{aligned} \widehat{\theta}_n : D_{\mathbf{X}}^n &\rightarrow \Theta \\ \mathbf{x} = (x_1, \dots, x_n) &\mapsto \widehat{\theta}_n(x_1, \dots, x_n) \end{aligned}$$

telle que $\forall \theta \in \Theta, L(x_1, \dots, x_n | \widehat{\theta}_n) \geq L(x_1, \dots, x_n | \theta)$.



1. $L(\mathbf{x}|\theta)$ n'a aucune raison d'être dérivable en θ .
2. Il n'y a aucune raison pour qu'un estimateur du maximum de vraisemblance soit sans biais.
3. Un estimateur du maximum de vraisemblance n'a aucune raison d'être unique.



Point méthode

1. Le principe de vraisemblance, à la base de la procédure d'estimation du maximum de vraisemblance, revient à rechercher la valeur de θ , fonction des observations $(x_1, \dots, x_n)$, qui assure la plus grande probabilité d'obtenir ces observations.
2. La recherche d'un maximum de la fonction de vraisemblance n'est pas forcément réduite à un simple calcul des zéros de la dérivée de L . Cependant, ce cas étant le plus fréquent, il est logique de poser les deux hypothèses suivantes :
 - $D_{\mathbf{X}}$ ne dépend pas de θ .
 - $\forall \mathbf{x} = (x_1, \dots, x_n) \in D_{\mathbf{X}}^n, \forall \theta \in \Theta, L$ est deux fois continûment dérivable par rapport à θ .
 Alors $\widehat{\theta}_n$, EMV, est solution du système mixte d'équation et d'inéquation en θ suivant :

Fiche 38

$$(I) \Leftrightarrow \begin{cases} \frac{dL}{d\theta}(\mathbf{x}|\theta) = 0 & (1) \\ \frac{d^2L}{d\theta^2}(\mathbf{x}|\theta) < 0. & (2) \end{cases}$$

3. En fait, vu la forme des densités des lois usuelles de probabilité, il est aussi aisé d'utiliser le logarithme de la vraisemblance, $\ln L(x_1, \dots, x_n | \theta)$, si $L(x_1, \dots, x_n | \theta) > 0$, pour tout $\mathbf{x} = (x_1, \dots, x_n) \in D_{\mathbf{X}}^n$, pour tout $\theta \in \Theta$, ce que nous supposons. Nous obtenons alors le système mixte d'équation et d'inéquation en θ suivant :

$$(1) \Leftrightarrow (1') \quad \frac{d \ln L}{d\theta}(\mathbf{x}|\theta) = 0$$

et

$$(2) \Leftrightarrow (2') \quad \frac{d^2 \ln L}{d\theta^2}(\mathbf{x}|\theta) < 0.$$



L'équation (1') s'appelle l'**équation de ln-vraisemblance**.

1. Estimation des paramètres d'une loi beta

Soit $(X_1, \dots, X_n)$ un échantillon aléatoire de loi parente la loi Beta($a; b$); $a > 0, b > 0$. Estimer $\theta = (a, b)$ par la méthode des moments. Nous rappelons qu'une variable aléatoire qui suit la loi beta, notée Beta($a; b$), admet pour fonction de densité, $f_X(x)$, qui est égale à :

$$\frac{1}{B(a; b)} x^{a-1} (1-x)^{b-1} \mathbb{1}_{]0;1[}(x).$$

2. Estimation des paramètres d'une loi uniforme continue

Soit $(X_1, \dots, X_n)$ un échantillon aléatoire de loi parente la loi uniforme $\mathcal{U}[a; b]$. Estimer $\theta = (a, b)$ par la méthode des moments.

3. Estimation du paramètre d'une loi de Poisson

Soit $(X_1, \dots, X_n)$ un échantillon aléatoire de loi parente la loi de Poisson $\mathcal{P}(\lambda)$. Estimer $\theta = \lambda$ par la méthode du maximum de vraisemblance.

4. Estimation du paramètre d'une loi uniforme continue

Soit $(X_1, \dots, X_n)$ un échantillon aléatoire de loi parente la loi uniforme $\mathcal{U}([0; \theta])$ avec $\theta > 0$. Estimer θ par la méthode du maximum de vraisemblance.

 Fiche 84

 Fiche 84

 Fiche 82

 Fiche 84

1. Soit X_1 de loi Beta($a; b$), nous avons :

$$\mathbb{E}_{(a,b)}(X_1) = \frac{a}{a+b} := c \quad \text{et} \quad \mathbb{E}_{(a,b)}(X_1^2) = \frac{a(a+1)}{(a+b)(a+b+1)} := d.$$

En inversant le système, nous obtenons, comme $d - c^2 = \text{Var}[X_1] > 0$:

$$a = \frac{c(c-d)}{d-c^2} \quad \text{et} \quad b = \frac{(1-c)(c-d)}{d-c^2}.$$

À l'aide de la loi forte des grands nombres, nous construisons des estimateurs de c et d :

$$\widehat{c}_n = \frac{1}{n} \sum_{i=1}^n X_i \quad \text{et} \quad \widehat{d}_n = \frac{1}{n} \sum_{i=1}^n X_i^2.$$

Nous déduisons de la méthode de substitution que :

$$\widehat{a}_n = \frac{\widehat{c}_n(\widehat{c}_n - \widehat{d}_n)}{\widehat{d}_n - \widehat{c}_n^2} \quad \text{et} \quad \widehat{b}_n = \frac{(1 - \widehat{c}_n)(\widehat{c}_n - \widehat{d}_n)}{\widehat{d}_n - \widehat{c}_n^2}$$

sont des estimateurs convergents de a et b .

2. Soit X_1 de loi $\mathcal{U}(a; b)$, nous avons :

$$\mathbb{E}_{(a,b)}(X_1) = \frac{a+b}{2} \quad \text{et} \quad \mathbb{E}_{(a,b)}(X_1^2) = \frac{a^2 + b^2 + ab}{3}$$

d'où

$$\frac{\widetilde{a}_n + \widetilde{b}_n}{2} = \frac{1}{n} \sum_{i=1}^n X_i = \bar{X}_n \quad \text{et} \quad \frac{\widetilde{a}_n^2 + \widetilde{b}_n^2 + \widetilde{a}_n \widetilde{b}_n}{3} = \frac{1}{n} \sum_{i=1}^n X_i^2.$$

Nous en déduisons les deux estimateurs convergents suivants :

$$\widetilde{a}_n = \bar{X}_n - \sqrt{\frac{3}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2} \quad \text{et} \quad \widetilde{b}_n = \bar{X}_n + \sqrt{\frac{3}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2}.$$

 Fiche 92

3. Nous avons :

$$L(x_1, \dots, x_n | \lambda) = \prod_{i=1}^n \frac{\lambda^{x_i} e^{-\lambda}}{x_i!}.$$

Nous obtenons :

$$\frac{\partial \ln L}{\partial \lambda}(\mathbf{x} | \lambda) = -n + \frac{S_n}{\lambda},$$

où

$$S_n = \sum_{i=1}^n x_i$$

et

$$\frac{\partial^2 \ln L}{\partial \lambda^2}(\mathbf{x} | \lambda) = -\frac{S_n}{\lambda^2}.$$

Ainsi, nous en déduisons l'estimateur du maximum de vraisemblance $\widehat{\lambda}_n = \frac{S_n}{n}$ et nous vérifions que :

$$\frac{\partial^2 \ln L}{\partial \lambda^2}(\mathbf{x} | \widehat{\lambda}_n) = -\frac{n^2}{S_n} < 0.$$

4. La vraisemblance associée à l'échantillon est définie par :

$$L_n(\mathbf{x} | \theta) = \prod_{i=1}^n \frac{1}{\theta} \mathbf{1}_{[0; \theta]}(x_i).$$

Déterminons l'estimateur du maximum de vraisemblance : la fonction $L_n(\mathbf{x} | \theta)$ atteint son maximum en θ si et seulement si :

$$\prod_{i=1}^n \mathbf{1}_{[0; \theta]}(x_i) = 1$$

car sinon elle est égale à 0. Autrement dit $\forall i = 1, \dots, n \quad \theta \geq x_i$. Donc, $0 \leq \sup_{1 \leq i \leq n} x_i \leq \theta$. Puis l'application $] \sup_{1 \leq i \leq n} x_i; +\infty[\mapsto 1/\theta \in \mathbb{R}_+^*$ est décroissante et ainsi l'estimateur du maximum de vraisemblance est égal à :

$$\widehat{\theta}_n = \sup_{1 \leq i \leq n} X_i.$$

Pour s'entraîner (solutions p. 553)

Fiche 81

108.1 Estimer par la méthode du maximum de vraisemblance et la méthode des moments le paramètre p d'une loi binomiale.

108.2 Estimer par la méthode du maximum de vraisemblance et la méthode des moments le paramètre μ d'une loi normale quand le paramètre σ^2 est connu.

Fiche 83

Exemples d'estimateurs de caractéristiques de position

Soit $(X_1, \dots, X_n)$ un échantillon aléatoire de loi parente la loi $\mathbb{P}_\theta$.

1. Estimateur de la moyenne

Définition 109.1

Un estimateur, noté $\widehat{\mu}_n$ ou parfois $\overline{X}_n$, de la moyenne μ est égal à :

$$\widehat{\mu}_n = \frac{1}{n} \sum_{i=1}^n X_i.$$

Théorème 109.1

1. Pour un échantillon aléatoire dont la loi parente admet une espérance notée μ , $\widehat{\mu}_n$ est un **estimateur sans biais** de la moyenne μ , c'est-à-dire $\mathbb{E}_\theta(\widehat{\mu}_n) = \mu$.
2. Lorsque la loi parente admet une variance, notée σ^2 , la variance de l'estimateur $\widehat{\mu}_n$ est égale à $\text{Var}_\theta(\widehat{\mu}_n) = \sigma^2/n$ et $\widehat{\mu}_n$ est un **estimateur convergent** de la moyenne μ .

 Fiche 106

2. Estimateur de la proportion

Dans le cas de l'estimateur d'une proportion π_A , les variables X_i prennent soit la valeur 0 avec une probabilité égale à $1 - \pi_A$ soit la valeur 1 avec une probabilité égale à π_A .

Définition 109.2

Un estimateur, noté $\widehat{\pi}_{A,n}$, de la proportion π_A est alors égal à :

$$\widehat{\pi}_{A,n} = \frac{1}{n} \sum_{i=1}^n X_i.$$

Théorème 109.2

1. $\widehat{\pi}_{A,n}$ est un **estimateur sans biais** de la proportion π_A c'est-à-dire $\mathbb{E}_\theta(\widehat{\pi}_{A,n}) = \pi_A$.
2. La variance de l'estimateur $\widehat{\pi}_{A,n}$ est égale à $\text{Var}_\theta(\widehat{\pi}_{A,n}) = \pi(1 - \pi)/n$ et $\widehat{\pi}_{A,n}$ est un **estimateur convergent** de la proportion π_A .



La proportion π_A est en fait une moyenne de 0 et de 1, ce qui explique les propriétés de l'estimateur $\widehat{\pi}_{A,n}$ qui est lui-même une moyenne.

3. Estimateur du minimum et du maximum

Soit $(X_1, \dots, X_n)$ un échantillon aléatoire. Pour chaque $\omega \in \Omega$, nous avons une valeur observée : $(X_1(\omega), \dots, X_n(\omega))$ de cet échantillon. Ordonnons cette suite de valeurs par ordre croissant :

$$X_{k_1(\omega)}(\omega) \leq X_{k_2(\omega)}(\omega) \leq \dots \leq X_{k_n(\omega)}(\omega),$$

où $k_1(\omega)$ est le numéro de la plus petite valeur, ..., $k_n(\omega)$ est le numéro de la plus grande valeur. Nous notons $X_{(i)}(\omega) = X_{k_i(\omega)}(\omega)$, la i -ième valeur rangée et le vecteur $(X_{(1)}, \dots, X_{(n)})$ s'appelle **l'échantillon rangé**.

Définition 109.3

Pour tout $i = 1, \dots, n$, la variable aléatoire $X_{(i)}$ est appelée la i -ième **statistique d'ordre de l'échantillon**.

 Fiche 100

Définition 109.4

$X_{(1)}$ est un estimateur du minimum et $X_{(n)}$ est un estimateur du maximum.

 Fiche 100

4. Estimateur de la médiane

Définition 109.5

Soient m et d les parties entière et décimale de $(n+1)/2$. Un **estimateur de la médiane**, $\widehat{Q}_2$ ou $\widehat{Q}_{0,5}$, est défini par :

 Fiche 100

$$\widehat{Q}_2 = X_{(m)} + d(X_{(m+1)} - X_{(m)}).$$



1. Il existe beaucoup de définitions pour l'estimateur de la médiane. Celle que nous avons décidé d'adopter ne permet pas de discussion au niveau du calcul et est utilisée par la plupart des logiciels de statistique.

2. Concrètement la formule indique que :

(a) si l'effectif de l'échantillon n est impair, $n = 2m + 1$ alors l'estimation de la médiane est égale à la valeur observée qui se situe au milieu de l'échantillon lorsque les valeurs de l'échantillon sont rangées par ordre croissant :

$$\widehat{Q}_2 = x_{(m)};$$

(b) si l'effectif de l'échantillon n est pair, $n = 2m$ alors l'estimation de la médiane est égale à la moyenne des valeurs observées qui se situent au centre de l'échantillon lorsque les valeurs de l'échantillon sont rangées par ordre croissant :

$$\widehat{Q}_2 = (x_{(m)} + x_{(m+1)})/2.$$

5. Estimateurs du premier et troisième quartiles

Définition 109.6

Soient m et d les parties entière et décimale de $(n+1)/4$. Un **estimateur du premier quartile**, $\widehat{Q}_1$ ou $\widehat{Q}_{0,25}$, est défini par :

 Fiche 100

$$\widehat{Q}_1 = X_{(m)} + d(X_{(m+1)} - X_{(m)}).$$

Définition 109.7

Soient m' et d' les parties entière et décimale de $3(n+1)/4$. Un **estimateur du troisième quartile**, $\widehat{Q}_3$ ou $\widehat{Q}_{0,75}$, est défini par :

 Fiche 100

$$\widehat{Q}_3 = X_{(m')} + d'(X_{(m'+1)} - X_{(m')}).$$

Albert A. Michelson et Edward Morley ont mesuré la vitesse de la lumière dans l'air entre le 5 juin 1879 et le 2 juillet 1879. Ils ont réalisé cinq expériences comportant chacune vingt mesures. La réponse expérimentale est la vitesse de la lumière dans l'air exprimée en km/s à laquelle il a été soustrait la valeur de 299000 km/s. Les résultats obtenus ont été reportés dans le tableau ci-après. La valeur de la vitesse de la lumière dans l'air communément acceptée de nos jours est de 299734,5 km/s.

850	740	900	1070	930	850	950	980	980	880	1000	980	930	650	760	810
1000	1000	960	960	960	940	960	940	880	800	850	880	900	840	830	790
810	880	880	830	800	790	760	800	880	880	880	860	720	720	620	860
970	950	880	910	850	870	840	840	850	840	840	840	890	810	810	820
800	770	760	740	750	760	910	920	890	860	880	720	840	850	850	780
890	840	780	810	760	810	790	810	820	850	870	870	810	740	810	940
950	800	810	870												

Source : A.J. Weekes, *A Genstat Primer*, London : Edward Arnold, 1986.

1. Cette série de résultats constitue-t-elle une population ? Un échantillon ? Justifier votre choix.
2. Estimer le minimum et le maximum de la population à partir de cet échantillon.
3. Estimer la moyenne de la population à partir de cet échantillon.
4. Estimer la médiane de la population à partir de cet échantillon.
5. Estimer le premier quartile et le troisième quartile de la population à partir de cet échantillon.

1. Il s'agit de la population des mesures réalisées par Michelson et Morley entre le 5 juin 1879 et le 2 juillet 1879. Par contre, cette série de résultats constitue un échantillon de toutes les mesures de la vitesse de la lumière qui ont été réalisées. C'est ainsi que nous les considérerons dans la suite de cette application.

2. Le minimum de l'échantillon, qui est égal à $x_{(1)} = \widehat{\min}(X)(obs) = 299620$ km/s, est une estimation du minimum de la population. Le maximum de l'échantillon, qui est égal à $x_{(100)} = \widehat{\max}(X)(obs) = 300070$ km/s, est une estimation du maximum de la population.

3. L'estimation de la moyenne de la population est égale à la moyenne de l'échantillon $\widehat{\mu}_n(obs) = 299852,40$ km/s.

4. La taille de l'échantillon est égale à 100 et est donc paire. Par conséquent, l'estimation de la médiane est égale à la moyenne des valeurs observées qui se situent au centre de l'échantillon, une fois que ces valeurs ont été rangées par ordre croissant. Nous obtenons comme estimation de la médiane : $\widehat{Q}_2(obs) = 299850,00$ km/s.

5. En appliquant les définitions présentées ci-dessus, nous obtenons les valeurs suivantes :

- l'estimation du premier quartile est égale à $\widehat{Q}_1(obs) = 299802,50$ km/s ;
- l'estimation du troisième quartile est égale à $\widehat{Q}_3(obs) = 299897,50$ km/s.

Un biologiste étudie un type d'algue qui attaque les plantes marines. La toxine contenue dans cette algue est obtenue sous forme d'une solution organique. Il mesure la quantité de toxine par gramme de solution. Il a obtenu les neuf mesures suivantes, exprimées en milligrammes : 1,2 ; 0,8 ; 0,6 ; 1,1 ; 1,2 ; 0,9 ; 1,5 ; 0,9 ; 1,0.

1. Estimer le minimum et le maximum de la population à partir de cet échantillon.
2. Estimer la moyenne de la population à partir de cet échantillon.
3. Estimer la médiane de la population à partir de cet échantillon.
4. Estimer le premier quartile et le troisième quartile de la population à partir de cet échantillon.

1. Le minimum de l'échantillon, qui est égal à $x_{(1)} = \min(X) = 0,60$ mg, est une estimation du minimum de la population. Le maximum de l'échantillon, qui est égal à $x_{(9)} = \max(X) = 1,50$ mg, est une estimation du maximum de la population.

2. L'estimation de la moyenne de la population est égale à la moyenne de l'échantillon $\hat{\mu}_n(obs) = 1,02$ mg.

3. La taille de l'échantillon est égale à 9 et est donc impaire. Par conséquent, l'estimation de la médiane est égale à la valeur au centre de l'échantillon, une fois que ces valeurs ont été rangées par ordre croissant. Nous obtenons comme estimation de la médiane : $\hat{Q}_2(obs) = 1,00$ mg.

4. En appliquant les définitions présentées ci-dessus, nous obtenons les valeurs suivantes :

- l'estimation du premier quartile est égale à $\hat{Q}_1(obs) = 0,85$ mg ;
- l'estimation du troisième quartile est égale à $\hat{Q}_3(obs) = 1,20$ mg.

Pour s'entraîner (solutions p. 553)

109.1 Émissions radioactives

L'appareil de mesures utilisé est un Compteur de Radioactivité Alpha et Bêta. Une source de Césium 137 a été placée à quelques centimètres du détecteur. Nous avons compté cinquante fois le nombre d'émissions pendant une seconde et obtenu les résultats suivants :

((4; 1); (5; 2); (6; 5); (7; 6); (8; 9); (9; 3); (10; 7); (11; 5); (12; 1); (13; 5); (14; 1); (15; 4); (17; 1)), où (4; 1) signifie qu'il a été observé une fois un décompte de quatre émissions pendant une seconde.

1. Estimer le minimum et le maximum de la population à partir de cet échantillon.

2. Estimer la moyenne de la population à partir de cet échantillon.

3. Estimer la médiane de la population à partir de cet échantillon.

4. Estimer le premier quartile et le troisième quartile de la population à partir de cet échantillon.

Exemples d'estimateurs de caractéristiques de dispersion

1. Estimateur de l'étendue

Définition 110.1

Un **estimateur**, noté $\widehat{e}_n$, de l'**étendue** $e(X)$ est égal à la différence entre la plus grande et la plus petite des valeurs prises par l'échantillon ou encore égal à :

$$\widehat{e}_n = X_{(n)} - X_{(1)},$$

où $X_{(n)}$ et $X_{(1)}$ ont été définis à la Fiche 109.

 Fiche 101

2. Intervalle interquartile et estimateur de l'étendue interquartile

Définition 110.2

La réalisation sur chaque échantillon de l'intervalle aléatoire $[\widehat{Q}_1; \widehat{Q}_3]$, appelé **intervalle interquartile de l'échantillon** et où $\widehat{Q}_1$ et $\widehat{Q}_3$ ont été définis à la Fiche 109, contient, sauf cas pathologique, 50% des observations de l'échantillon.

 Fiche 101

Définition 110.3

Un **estimateur**, noté $\widehat{EIQ}_n$, de l'**étendue interquartile** $EIQ(X)$, est égal à la longueur de l'intervalle interquartile c'est-à-dire à la différence entre l'estimateur du troisième quartile, $\widehat{Q}_3$, et l'estimateur du premier quartile, $\widehat{Q}_1$:

$$\widehat{EIQ}_n = \widehat{Q}_3 - \widehat{Q}_1.$$

 Fiche 101

3. Estimateurs de la variance

► Estimateur biaisé de la variance

Définition 110.4

Un **estimateur**, noté S_n^2 et souvent simplement S^2 , de la **variance** σ^2 est égal à :

$$S_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \widehat{\mu}_n)^2.$$

 Fiche 101

Théorème 110.1

1. Pour un échantillon aléatoire dont la loi parente admet une espérance notée μ et une variance notée σ^2 , S_n^2 est un **estimateur biaisé** de la variance σ^2 .
2. Le biais de cet estimateur, noté $B(S_n^2)$, est égal à $-\sigma^2/n$.
3. S_n^2 est un **estimateur convergent et asymptotiquement sans biais**.

 Fiche 106



La variance de cet estimateur est égale à :

$$\text{Var}_\theta(S_n^2) = \frac{n-1}{n^3} \left[(n-1)_\mu m_4(X) - (n-3)\sigma^4 \right],$$

où $_\mu m_4(X)$ désigne le moment centré d'ordre 4. Ce résultat qui semble à première vue compliqué est obtenu après une série « un peu longue » de calculs.

Fiche 101

► Estimateur sans biais de la variance

Définition 110.5

Un **estimateur corrigé**, noté $S_{n,c}^2$ et souvent simplement S_c^2 , de la variance σ^2 est égal à :

$$S_{n,c}^2 = \frac{nS_n^2}{n-1} = \frac{1}{n-1} \sum_{i=1}^n (X_i - \widehat{\mu}_n)^2.$$

Théorème 110.2

1. Pour un échantillon aléatoire dont la loi parente admet une espérance notée μ et une variance notée σ^2 , $S_{n,c}^2$ est un **estimateur sans biais** de la variance σ^2 .
2. $S_{n,c}^2$ est un **estimateur convergent** pour la variance σ^2 .

4. Estimateurs de l'écart-type

► Estimateur de l'écart-type déduit de l'estimateur corrigé de la variance

Définition 110.6

Un **estimateur**, noté $\widehat{\sigma}_n = \sqrt{S_{n,c}^2}$, de l'écart-type σ est égal à :

$$\widehat{\sigma}_n = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \widehat{\mu}_n)^2}.$$

Fiche 101



L'estimateur $\widehat{\sigma}_n$ est également souvent noté $S_{n,c}$ ou simplement S_c .

Théorème 110.3

1. Pour un échantillon aléatoire dont la loi parente admet une espérance notée μ et une variance notée σ^2 , $\widehat{\sigma}_n$ est généralement un **estimateur biaisé** de l'écart-type σ .
2. Dans le cas d'une loi parente normale, le biais de cet estimateur, noté $B(\widehat{\sigma}_n)$, est égal à :

$$\mathbb{E}(\widehat{\sigma}_n) - \sigma = \sqrt{\frac{2}{n-1}} \frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)} \sigma - \sigma = \left(-\frac{1}{4n} + o\left(\frac{1}{n}\right) \right) \sigma,$$

où $o(1/n)$ est une fonction qui tend vers 0 plus vite que $1/n$ et Γ est la fonction Γ d'Euler.

Fiches 41 et 53

3. $\widehat{\sigma}_n$ est un *estimateur asymptotiquement sans biais*.



La racine carrée d'un estimateur sans biais n'est donc pas nécessairement un estimateur sans biais.

Fiche 106

► **Estimateur corrigé de l'écart-type**

Définition 110.7

Un **estimateur corrigé**, noté $\widehat{\sigma}_{n,\text{corrigé}}$, de l'écart-type σ est égal à :

$$\widehat{\sigma}_{n,\text{corrigé}} = \sqrt{\frac{n-1}{2}} \frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} \widehat{\sigma}_n \approx \sqrt{\frac{2n-2}{2n-3}} \widehat{\sigma}_n \approx \left(1 + \frac{1}{4n-6}\right) \widehat{\sigma}_n.$$

Théorème 110.4

Pour un échantillon aléatoire de loi parente une loi normale d'espérance μ et de variance σ^2 , $\widehat{\sigma}_{n,\text{corrigé}}$ est un *estimateur sans biais* de l'écart-type σ .

► **Utilisation de la médiane des écarts absolus à la médiane**

La **médiane des écarts absolus à la médiane** multipliée par $1/\Phi^{-1}(3/4) \approx 1,4826$, avec Φ^{-1} la fonction quantile de la loi normale centrée et réduite, permet d'estimer de manière convergente et robuste l'écart-type d'un caractère distribué selon une loi normale.

Fiche 101

5. Estimateur du coefficient de variation

Définition 110.8

Un **estimateur**, noté $\widehat{CV}_n$, du **coefficient de variation** $CV(X)$ est égal à :

$$\widehat{CV}_n = \frac{\widehat{\sigma}_n}{\widehat{\mu}_n}.$$

Fiche 101

Théorème 110.5

1. Pour un échantillon aléatoire dont la loi parente est une loi normale d'espérance μ et de variance σ^2 , $\widehat{CV}_n$ est un *estimateur biaisé* du coefficient de variation $CV(X)$.
2. Pour une série de données issue d'une loi normale et ne pouvant prendre des valeurs négatives qu'avec une probabilité très faible, un estimateur du coefficient de variation, sans biais jusqu'à l'ordre n^{-1} , pour un échantillon de taille n est égal à :

$$\widehat{CV}_n^* = \widehat{CV}_n - \frac{1}{4(n-1)} + \frac{1}{n} \widehat{CV}_n^2.$$

Reprendre les données correspondantes de la Fiche 109.

1. Calculer une estimation de l'étendue.
2. Calculer l'intervalle interquartile de l'échantillon puis une estimation de l'étendue interquartile.
3. Calculer une estimation biaisée et sans biais de la variance
4. Dans cette question, nous faisons l'hypothèse supplémentaire, dont nous vérifierons la validité à la Fiche 118, que les résultats de l'expérience sont des réalisations d'une variable aléatoire qui suit une loi normale. En déduire une estimation biaisée et sans biais de l'écart-type.
5. Calculer une estimation du coefficient de variation. En faisant la même hypothèse supplémentaire qu'à la Question 4., donner une estimation sans biais du coefficient de variation.

1. 450,00 km/s pour l'estimation de l'étendue.
2. [802,5 km/s; 897,5 km/s] pour l'intervalle interquartile de l'échantillon donc $897,5 - 802,5 = 95,00$ km/s pour l'estimation de l'étendue interquartile.
3. $6180,24 \text{ (km/s)}^2$ pour l'estimation de la variance avec S_n^2 donc biaisée, $6242,67 \text{ (km/s)}^2$ pour l'estimation de la variance avec $S_{n,c}^2$ donc sans biais.
4. 79,01 km/s pour l'estimation de l'écart-type avec $\hat{\sigma}_n$ donc biaisée, 79,21 km/s pour l'estimation de l'écart-type avec $\hat{\sigma}_{n,\text{corrigé}}$ donc sans biais.
5. L'estimation du coefficient de variation obtenue avec $\widehat{CV}_n$ est égale à 0,0263 %. Dans le contexte de normalité supposée, nous obtenons une estimation sans biais de 0,0264 % en utilisant $\widehat{CV}_n^*$.

Pour s'entraîner (solutions p. 554)

110.1 Émissions radioactives

En reprenant les données de l'exercice 109.1, répondre aux questions suivantes.

1. Calculer une estimation de l'étendue.
2. Calculer l'intervalle interquartile de l'échantillon puis une estimation de l'étendue interquartile.
3. Calculer une estimation biaisée et sans biais de la variance et de l'écart-type.
4. Calculer une estimation du coefficient de variation.

 Fiche 83

Estimation par intervalle de confiance et intervalle de confiance pour une proportion

1. Principe

Définition 111.1

Lorsque nous procédons à l'estimation de l'une des caractéristiques, notée θ , de la population, il est plus pertinent de fournir un intervalle $\theta_1 < \theta < \theta_2$ plutôt que de donner simplement une estimation ponctuelle, c'est-à-dire d'écrire uniquement $\widehat{\theta} = c$. Un tel intervalle $] \theta_1; \theta_2[$ s'appelle une **estimation par intervalle de confiance** du paramètre θ et est une **estimation ensembliste** de la caractéristique θ .

La méthode sur laquelle repose la construction des intervalles de confiance est la suivante :

- Soit $\widehat{\theta}_n$ un estimateur de θ dont nous connaissons la loi de probabilité pour chaque valeur de θ . Il va de soi qu'il convient d'utiliser le meilleur estimateur possible selon les critères de la Fiche 106.
- Étant donné une valeur θ_0 du paramètre θ , nous déterminons un **intervalle de probabilité bilatéral de niveau $(1 - \alpha)$** pour l'estimateur $\widehat{\theta}_n$, c'est-à-dire deux bornes θ_1^n et θ_2^n telles que :

$$\mathbb{P}_{\theta_0}(\theta_1^n < \widehat{\theta}_n < \theta_2^n) \geq 1 - \alpha.$$

- Nous réalisons cette opération pour toutes les valeurs θ_0 du paramètre θ .



1. Les deux bornes θ_1^n et θ_2^n dépendent de la valeur θ_0 .
2. Nous pourrions également construire des **intervalles unilatéraux** pour lesquels $\theta_1^n = -\infty$ ou $\theta_2^n = +\infty$.
3. Nous choisissons dans la plupart des cas un intervalle de probabilité bilatéral à risques symétriques $\alpha/2$ et $\alpha/2$.
4. Si nous augmentons le **niveau $1 - \alpha$** , nous augmentons la **longueur $\theta_2^n - \theta_1^n$** de l'intervalle de probabilité bilatéral.

Fiche 106

2. Intervalle de confiance

Définition 111.2

Soit $(X_1, \dots, X_n)$ un échantillon aléatoire de loi parente la loi $\mathbb{P}_\theta$, avec $\theta \in \Theta \subset \mathbb{R}$. Nous appelons **intervalle de confiance de niveau de confiance $1 - \alpha$ du paramètre θ** tout intervalle $] \theta_1(X_1, \dots, X_n); \theta_2(X_1, \dots, X_n)[$ tel que : $\mathbb{P}_\theta(\theta_1(X_1, \dots, X_n); \theta_2(X_1, \dots, X_n) \ni \theta) = \mathbb{P}_\theta(\theta_1(X_1, \dots, X_n) < \theta \text{ et } \theta_2(X_1, \dots, X_n) > \theta) = 1 - \alpha$ pour $\alpha \in [0; 1]$ fixé.



1. La probabilité α pour que l'intervalle de confiance ne contienne pas la vraie valeur θ peut être répartie différemment de part et d'autre des bornes de l'intervalle de confiance. Posons $\alpha = \alpha_1 + \alpha_2$, où α_1 mesure le risque que la vraie valeur inconnue de θ soit plus petite que la borne inférieure θ_1 de l'intervalle de confiance et α_2 mesure le risque que la vraie valeur inconnue de θ soit plus grande que la borne supérieure θ_2 de l'intervalle de confiance.
2. L'**intervalle de confiance** est dit **bilatéral** lorsque $\alpha_1 \neq 0$ et $\alpha_2 \neq 0$. Si $\alpha_1 = \alpha_2 = \alpha/2$, l'intervalle est dit symétrique, sinon il est dit dissymétrique.
3. L'**intervalle de confiance** est dit **unilatéral** lorsque $\alpha_1 = 0$ ou $\alpha_2 = 0$. Il est toujours dissymétrique.
 - Si nous sommes uniquement intéressés par le calcul d'une valeur minimale pour la quantité estimée, nous fixons $\alpha_1 = \alpha$ et $\alpha_2 = 0$. L'intervalle de confiance est alors de la forme $]\theta_1; \theta_2[=]a; +\infty[$.
 - Si nous sommes uniquement intéressés par le calcul d'une valeur maximale pour la quantité estimée, nous fixons $\alpha_1 = 0$ et $\alpha_2 = \alpha$. L'intervalle de confiance est alors de la forme $]\theta_1; \theta_2[=]-\infty; b[$.
4. $]\theta_1; \theta_2[$ est un intervalle aléatoire car il dépend de l'estimateur $\widehat{\theta}_n$. $]\theta_1; \theta_2[$ s'obtient par :

$$\begin{cases} \theta_1 = (\theta_2^n)^{-1}(\widehat{\theta}_n) \\ \theta_2 = (\theta_1^n)^{-1}(\widehat{\theta}_n). \end{cases}$$

5. Si nous augmentons le **niveau $1 - \alpha$** ou encore parfois appelé également le **coefficient de confiance**, nous augmentons la **longueur** $\theta_2 - \theta_1$ de l'intervalle de confiance.
6. Nous associons souvent aux intervalles de confiance, un test, basé sur la règle de décision suivante : soit $\widehat{\theta}(obs)$ la valeur observée de l'estimateur $\widehat{\theta}_n$:
 - si la valeur observée $\widehat{\theta}(obs) \in]\theta_1^n; \theta_2^n[$, nous conservons la valeur θ_0 comme valeur possible du paramètre θ ;
 - si la valeur observée $\widehat{\theta}(obs) \notin]\theta_1^n; \theta_2^n[$, nous éliminons la valeur θ_0 .

3. Intervalle de confiance pour une proportion

Nous souhaitons construire un intervalle de confiance pour une proportion π_A d'individus de la population qui possèdent un certain caractère A . Pour estimer π_A , nous allons nous servir de l'estimateur $\widehat{\pi}_{n,A}$ à partir du moment où nous faisons l'hypothèse que le tirage est aléatoire avec remise, ce qui correspond à une population infinie. D'autre part, nous pouvons montrer que $n\widehat{\pi}_{n,A}$ suit la loi binomiale de paramètres n et π_A . À partir de ce résultat, nous pouvons construire un intervalle de confiance pour π_A . Les trois méthodes pour construire un intervalle de confiance pour une proportion π_A que nous rencontrerons le plus souvent sont :

1. la **méthode exacte** ou encore appelée la **méthode de Clopper-Pearson** qui peut être mise en œuvre à l'aide du nomogramme de Larson ou avec la plupart des logiciels de statistique ;
2. la **méthode du score** ou encore appelée la **méthode de Wilson** ou encore la **méthode de l'ellipse** ;
3. la **méthode asymptotique** ou encore appelée la **méthode de Wald**.

 Fiche 81



Des études statistiques ont montré que, parmi ces trois méthodes, la **méthode à privilégier** est la **méthode du score**. Néanmoins, la méthode de Wald reste très utilisée et présentée dans de nombreux manuels alors qu'elle ne permet généralement pas d'obtenir des intervalles de confiance de qualité convenable. Ce défaut est encore plus marqué pour des valeurs de π proches de 0 ou 1.

Définition 111.3 Intervalle de confiance bilatéral par la méthode du score dite de Wilson

L'intervalle de confiance bilatéral pour la proportion π_A au niveau de confiance $(1 - \alpha)$ est égal à :

$$\frac{\widehat{\pi}_{n,A}(obs) + \frac{1}{2n}z_{1-\alpha/2}^2 - z_{1-\alpha/2} \times \sqrt{\frac{\widehat{\pi}_{n,A}(obs)(1 - \widehat{\pi}_{n,A}(obs))}{n} + \frac{z_{1-\alpha/2}^2}{4n^2}}}{1 + \frac{1}{n}z_{1-\alpha/2}^2} < \pi_A$$

$$< \frac{\widehat{\pi}_{n,A}(obs) + \frac{1}{2n}z_{1-\alpha/2}^2 + z_{1-\alpha/2} \times \sqrt{\frac{\widehat{\pi}_{n,A}(obs)(1 - \widehat{\pi}_{n,A}(obs))}{n} + \frac{z_{1-\alpha/2}^2}{4n^2}}}{1 + \frac{1}{n}z_{1-\alpha/2}^2},$$

où $z_{1-\alpha/2}$ est le quantile d'ordre $1 - \alpha/2$ pour la loi normale centrée et réduite.



Fiche 83

Définition 111.4 Intervalle de confiance bilatéral par la méthode asymptotique dite de Wald

L'intervalle de confiance bilatéral pour la proportion π_A au niveau de confiance $(1 - \alpha)$ est égal à :

$$\widehat{\pi}_{n,A}(obs) - z_{1-\alpha/2} \times \sqrt{\frac{\widehat{\pi}_{n,A}(obs)(1 - \widehat{\pi}_{n,A}(obs))}{n}} < \pi_A$$

$$< \widehat{\pi}_{n,A}(obs) + z_{1-\alpha/2} \times \sqrt{\frac{\widehat{\pi}_{n,A}(obs)(1 - \widehat{\pi}_{n,A}(obs))}{n}},$$

où $z_{1-\alpha/2}$ est le quantile d'ordre $1 - \alpha/2$ pour la loi normale centrée et réduite.



Fiche 83



1. Il existe pour ces deux intervalles une formule avec une **correction de continuité de Yates** pour tenir compte du passage d'une loi discrète à une loi continue. Donc, parfois, nous pouvons constater un léger écart entre le calcul de cette formule si nous le réalisons à la main et le calcul donné par un logiciel de statistique. D'où cette recommandation : regarder ce qui est programmé dans le manuel du logiciel utilisé.
2. Pour établir l'intervalle de confiance par la méthode de Wald, l'approximation de la loi binomiale par la loi normale a été utilisée. Il faut se rappeler que pour utiliser cette approximation, certaines conditions doivent être remplies.
3. Enfin, il manque la formule mathématique de l'intervalle de confiance par la méthode de Clopper-Pearson. Elle n'a pas été présentée ici car le calcul est un peu plus compliqué que les deux autres et fait partie des compléments en ligne.



Fiche 94

Dans le *Ouest-France* du samedi 23 janvier 2010, vous pouvez lire : « Plus de garçons que de filles ! Avec 507 bébés mâles comptabilisés à Saint-Lô en 2009, contre 481 fillettes, les naissances masculines sont toujours plus nombreuses. » Notons π la proportion de garçons.

1. Construire un intervalle de confiance bilatéral de niveau 95 % autour de la proportion de garçons π avec la méthode de Clopper-Pearson.
2. Même question avec la méthode du score.
3. Même question avec la méthode de Wald.
4. Pour une famille de Saint-Lô et avec une confiance de 95 %, faut-il attribuer le déséquilibre observé en 2009 entre la probabilité d'avoir une fille ou celle d'avoir un garçon au simple hasard, $\pi = 0,5$, ou conclure à une différence réelle $\pi \neq 0,5$?
5. À partir de quelle taille d'échantillon n_0 , un tel déséquilibre entre les naissances des garçons et des filles n'aurait-il pas pu être lié, avec une confiance de 95 %, au simple hasard ?

1. Obtenue en utilisant le langage **R** ou le nomogramme de Larson, la réalisation sur l'échantillon de l'intervalle de confiance bilatéral de niveau 95 % autour de la proportion de garçons π avec la méthode de Clopper-Pearson est $[0,481; 0,545]$.

2. La réalisation sur l'échantillon de l'intervalle de confiance bilatéral de niveau 95 % autour de la proportion de garçons π avec la méthode du score est $[0,482; 0,544]$.

3. La réalisation sur l'échantillon de l'intervalle de confiance bilatéral de niveau 95 % autour de la proportion de garçons π avec la méthode de Wald est $[0,482; 0,544]$.

4. La taille de l'échantillon est égale à 988 et est élevée : les réalisations des trois intervalles de confiance bilatéraux sont très proches. Nous constatons que, bien que l'estimation ponctuelle de la proportion π à partir de l'échantillon soit égale à $\widehat{\pi}(obs) = 0,513$, chacune des réalisations des intervalles de confiance bilatéraux contient la valeur 0,5. Par conséquent, il n'est pas possible d'exclure le fait que le déséquilibre entre le nombre de garçons et de filles observé en 2009 à Saint-Lô ne soit dû qu'aux caprices du hasard à partir de la situation où la proportion π serait égale à 0,5.

5. L'effectif recherché est supérieur à 988, celui de l'échantillon qui vient d'être étudié. Par conséquent, les réalisations des trois intervalles de confiance bilatéraux seront quasiment confondues. En utilisant la méthode du score, nous obtenons $n_0 = 5544$.

 Fiche 121

Pour s'entraîner (solutions p. 554)

111.1 Une machine fabrique des billes métalliques. Un client achète un lot de billes qui est censé avoir été fabriqué selon un cahier des charges qui avait accepté par le fournisseur. Un contrôle de qualité est mené sur ce lot afin de déterminer s'il est conforme aux spécifications qui avaient été fixées. Il est procédé au contrôle de 100 billes choisies

au hasard et avec remise dans le lot. Quatre d'entre elles sont non conformes. La proportion maximale de non conformes tolérée par le client est de 0,5 %. Compte tenu des résultats observés sur cet échantillon, le client doit-il, avec une confiance de 95 %, accepter ce lot ou le renvoyer au fournisseur ?

Intervalles de confiance pour une espérance et une variance

Théorème 112.1 Théorème de Fisher

Soit $(X_1, \dots, X_n)$ un échantillon aléatoire de loi parente la loi normale d'espérance μ et d'écart-type σ .

- $\widehat{\mu}_n$ et $nS_n^2/\sigma^2 = (n-1)S_{n,c}^2/\sigma^2$ sont deux variables aléatoires indépendantes, S_n^2 et $S_{n,c}^2$ ont été définis à la Fiche 110.
- $\frac{\widehat{\mu}_n - \mu}{\sigma/\sqrt{n}}$ suit la loi normale $\mathcal{N}(0; 1)$ et $nS_n^2/\sigma^2 = (n-1)S_{n,c}^2/\sigma^2$ suit la loi du Khi-deux $\chi^2(n-1)$.
- $\widehat{\mu}_n$ est le meilleur estimateur sans biais de μ .
- $\frac{\widehat{\mu}_n - \mu}{S_{n,c}/\sqrt{n}} = \frac{\widehat{\mu}_n - \mu}{S_n/\sqrt{n-1}}$ suit la loi de Student $t(n-1)$.
- $\widehat{\sigma}_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2$ est un estimateur de σ^2 et $n\widehat{\sigma}_n^2/\sigma^2$ suit la loi du Khi-deux $\chi^2(n)$.
- S_n^2 est un estimateur de σ^2 biaisé et $S_{n,c}^2$ est un estimateur sans biais de σ^2 .

Fiche 110

Fiche 85

Fiche 85

Fiche 110

1. Espérance d'une loi normale, écart-type σ connu

Définition 112.1

L'intervalle de probabilité bilatéral pour $\widehat{\mu}_n$ à $1 - \alpha$ est égal à :

$$\mu - z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}} < \widehat{\mu}_n < \mu + z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}}.$$

Définition 112.2

L'intervalle de confiance bilatéral pour μ à $1 - \alpha$ est égal à :

$$\widehat{\mu}_n - z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}} < \mu < \widehat{\mu}_n + z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}}.$$



- $z_{1-\alpha/2}$ désigne le quantile d'ordre $1 - \alpha/2$ pour la loi normale centrée et réduite.
- Souvent $1 - \alpha$ sera fixé à 0,95 et par conséquent $z_{1-\alpha/2}$ sera égal à 1,96.
- Pour chaque réalisation, $\widehat{\mu}_n(obs)$, de $\widehat{\mu}_n$ calculée à partir de l'observation d'un échantillon de taille n , nous obtenons une réalisation de l'intervalle de confiance.

2. Espérance d'une loi normale, écart-type σ inconnu

Définition 112.3

L'intervalle de probabilité bilatéral pour $\frac{\widehat{\mu}_n - \mu}{S_{n,c}/\sqrt{n}}$ à $1 - \alpha$ est égal à :

$$-t_{n-1; 1-\alpha/2} < \frac{\widehat{\mu}_n - \mu}{S_{n,c}/\sqrt{n}} < t_{n-1; 1-\alpha/2}.$$

Définition 112.4

L'**intervalle de confiance** bilatéral pour μ à $1 - \alpha$ est égal à :

$$\widehat{\mu}_n - t_{n-1;1-\alpha/2} \frac{S_{n,c}}{\sqrt{n}} < \mu < \widehat{\mu}_n + t_{n-1;1-\alpha/2} \frac{S_{n,c}}{\sqrt{n}}.$$



- $t_{n-1;1-\alpha/2}$ est le quantile d'ordre $1 - \alpha/2$ pour la loi de Student $t(n - 1)$.
- $S_n/S_{n,c} = \sqrt{n-1}/\sqrt{n}$.
- Pour chaque réalisation, $(\widehat{\mu}_n(obs), S_{n,c}(obs))$, du couple $(\widehat{\mu}_n, S_{n,c})$ calculée à partir de l'observation d'un échantillon de taille n , nous obtenons une réalisation de l'intervalle de confiance.

3. Variance d'une loi normale, espérance μ connue**Définition 112.5**

k_1 et k_2 sont les bornes d'un **intervalle de probabilité** bilatéral pour $n\widehat{\sigma}_n^2/\sigma^2$ si

$$\mathbb{P}\left(k_1 < n\widehat{\sigma}_n^2/\sigma^2 < k_2\right) = 1 - \alpha.$$

Par exemple, nous pouvons prendre k_1 égal au quantile d'ordre $\alpha/2$ pour la loi du Khi-deux $\chi^2(n)$ et k_2 égal au quantile d'ordre $1 - \alpha/2$ pour la loi du Khi-deux $\chi^2(n)$.

Définition 112.6

Un **intervalle de confiance** bilatéral pour σ^2 à $1 - \alpha$ est égal à :

$$n\widehat{\sigma}_n^2/k_2 < \sigma^2 < n\widehat{\sigma}_n^2/k_1.$$



Pour chaque réalisation, $\widehat{\sigma}_n^2(obs)$, de $\widehat{\sigma}_n^2$ calculée à partir de l'observation d'un échantillon de taille n , nous obtenons une réalisation de l'intervalle de confiance.

4. Variance d'une loi normale, espérance μ inconnue**Définition 112.7**

l_1 et l_2 sont les bornes d'un **intervalle de probabilité** bilatéral pour $(n - 1)S_{n,c}^2/\sigma^2$ si

$$\mathbb{P}\left(l_1 < nS_{n,c}^2/\sigma^2 = (n - 1)S_{n,c}^2/\sigma^2 < l_2\right) = 1 - \alpha.$$

Par exemple, nous pouvons prendre l_1 égal au quantile d'ordre $\alpha/2$ pour la loi du Khi-deux $\chi^2(n - 1)$ et l_2 égal au quantile d'ordre $1 - \alpha/2$ pour la loi du Khi-deux $\chi^2(n - 1)$.

Définition 112.8

Un **intervalle de confiance** bilatéral pour σ^2 à $1 - \alpha$ est égal à :

$$(n - 1)S_{n,c}^2/l_2 = nS_{n,c}^2/l_2 < \sigma^2 < nS_{n,c}^2/l_1 = (n - 1)S_{n,c}^2/l_1.$$



Pour chaque réalisation, $S_{n,c}(obs)$, de $S_{n,c}$ calculée à partir de l'observation d'un échantillon de taille n , nous obtenons une réalisation de l'intervalle de confiance.

1. Détermination de la vitesse de la lumière

Reprendre les données correspondantes de la Fiche 109 en supposant qu'elles sont la réalisation d'un échantillon aléatoire de taille 100 de loi parente la loi normale d'espérance μ et d'écart-type σ .

a. Déterminer la réalisation sur cet échantillon de l'intervalle de confiance bilatéral à 95% pour l'espérance μ de la vitesse de la lumière dans l'air.

b. Déterminer la réalisation sur cet échantillon de l'intervalle de confiance bilatéral à 95% pour la variance σ^2 de la vitesse de la lumière dans l'air.

2. Les plantes marines

Reprendre les données correspondantes de la Fiche 109 en supposant qu'elles sont la réalisation d'un échantillon aléatoire de taille 9 de loi parente la loi normale d'espérance μ et d'écart-type σ .

a. Déterminer la réalisation sur cet échantillon de l'intervalle de confiance bilatéral à 95% pour l'espérance μ de la quantité de toxine par gramme de solution.

b. Déterminer la réalisation sur cet échantillon de l'intervalle de confiance bilatéral à 95% pour la variance σ^2 de la quantité de toxine par gramme de solution.

1.a. Déterminons la réalisation de l'intervalle de confiance bilatéral à 95% pour l'espérance μ de la vitesse de la lumière dans l'air. L'espérance μ et l'écart-type σ étant inconnus, la réalisation de l'intervalle de confiance bilatéral pour μ de niveau 95% s'obtient avec la formule suivante :

$$\widehat{\mu}_{100}(obs) - t_{99;0,975} \frac{S_{100,c}(obs)}{\sqrt{100}} < \mu < \widehat{\mu}_{100}(obs) + t_{99;0,975} \frac{S_{100,c}(obs)}{\sqrt{100}}$$

$$299836,7 \text{ km/s} < \mu < 299868,1 \text{ km/s}$$

où $t_{99;0,975} = 1,984217$ est le quantile d'ordre 0,975 pour la loi de Student à 99 degrés de liberté $t(99)$, commande `qt(.975, 99)` avec le langage **R**.

b. Déterminons la réalisation de l'intervalle de confiance bilatéral à 95% pour la variance σ^2 de la vitesse de la lumière dans l'air. L'espérance μ et l'écart-type σ étant inconnus, une réalisation de l'intervalle de confiance bilatéral pour σ^2 de niveau 95% s'obtient avec la formule suivante :

$$99S_{100,c}^2(obs)/l_2 < \sigma^2 < 99S_{100,c}^2(obs)/l_1$$

$$4812,447 \text{ (km/s)}^2 < \sigma^2 < 8424,412 \text{ (km/s)}^2$$

où $l_1 = 73,36108$ est le quantile d'ordre 0,025 pour la loi du Khi-deux $\chi^2(99)$ et $l_2 = 128,42199$ est le quantile d'ordre 0,975 pour la loi du Khi-deux $\chi^2(99)$, commandes `qchisq(.025, 99)` et `qchisq(.975, 99)` avec le langage **R**.

2.a. Déterminons la réalisation de l'intervalle de confiance bilatéral à 95 % pour l'espérance μ de la quantité de toxine par gramme de solution. L'espérance μ et l'écart-type σ étant inconnus, une réalisation de l'intervalle de confiance pour μ de niveau 95% s'obtient avec la formule suivante :

$$\widehat{\mu}_9(obs) - t_{8;0,975} \frac{S_{9,c}(obs)}{\sqrt{9}} < \mu < \widehat{\mu}_9(obs) + t_{8;0,975} \frac{S_{9,c}(obs)}{\sqrt{9}}$$

$$0,820 \text{ mg} < \mu < 1,225 \text{ mg}$$



Fiche 83



Fiche 83



Fiche 85



Fiche 85

où $t_{8,0,975} = 2,3060$ est le quantile d'ordre 0,975 pour la loi de Student à huit degrés de liberté $t(8)$, commande `qt(.975, 8)` avec le langage **R**.

b. Déterminons la réalisation de l'intervalle de confiance bilatéral à 95 % pour la variance σ^2 de la quantité de toxine par gramme de solution. L'espérance μ et l'écart-type σ étant inconnus, une réalisation de l'intervalle de confiance pour σ^2 de niveau 95 % s'obtient avec la formule suivante :

$$8S_{9,c}^2(obs)/l_2 < \sigma^2 < 8S_{9,c}^2(obs)/l_1$$

$$0,032 \text{ mg}^2 < \sigma^2 < 0,255 \text{ mg}^2$$

où $l_1 = 2,1797$ est le quantile d'ordre 0,025 pour la loi du Khi-deux $\chi^2(8)$ et $l_2 = 17,5345$ est le quantile d'ordre 0,975 pour la loi du Khi-deux $\chi^2(8)$, commandes `qchisq(.025, 8)` et `qchisq(.975, 8)` avec le langage **R**.

Fiche 85

Fiche 85

Pour s'entraîner (solutions p. 554)

112.1 Voulant évaluer rapidement les résultats obtenus par ses 200 élèves ingénieurs lors d'un examen, un professeur décide de corriger quelques copies tirées au hasard. Il admet par ailleurs que les notes de ses élèves sont des réalisations d'une loi normale de variance 4.

1. Le professeur corrige un échantillon de 7 copies et trouve une moyenne de 11. Quelle est la réalisation de l'intervalle de confiance bilatéral à 95 % de l'espérance obtenue à partir des 200 copies ?
2. Combien de copies le professeur doit-il corriger s'il veut situer l'espérance de la note d'un élève dans un intervalle de confiance d'amplitude 2, avec un risque de 5 % ?
3. En trouvant une moyenne égale à 11, combien de copies le professeur devrait-il corriger pour pouvoir dire, avec un risque de 1 %, que la moyenne de tous les élèves est supérieure à 10 ?

112.2

www

Une entreprise fabrique un certain type de composants électroniques dont la durée de vie X , exprimée en heures, est une variable aléatoire. Des mesures effectuées sur un échantillon aléatoire de taille 50 ont donné les résultats suivants :

$$\sum_{i=1}^{50} x_i = 60000; \quad \sum_{i=1}^{50} x_i^2 = 74 \times 10^6.$$

1. Donner une estimation ponctuelle de la durée de vie moyenne des composants.
2. Donner une estimation ponctuelle de l'écart-type de cette durée de vie.
3. Donner une réalisation de l'intervalle de confiance bilatéral à 95 %, puis à 99 % de cette durée de vie moyenne.
4. Quelle aurait dû être la taille de l'échantillon pour qu'une réalisation de l'intervalle de confiance bilatéral à 95 % de la durée de vie moyenne des composants ait une amplitude de 60 heures ?

Introduction à la théorie des tests d'hypothèses

1. Introduction à la notion de tests

Définition 113.1

Un **test** est un mécanisme qui permet de trancher entre deux hypothèses à la vue des résultats d'un échantillon, en quantifiant le risque associé à la décision prise.

Définition 113.2

Soit $\mathcal{H}_0$ et $\mathcal{H}_1$ deux hypothèses, dont une et une seule est vraie. $\mathcal{H}_0$ joue le plus souvent un rôle prédominant par rapport à $\mathcal{H}_1$. Cela est la conséquence du fait que $\mathcal{H}_0$ est l'**hypothèse de référence**, aussi appelée **hypothèse nulle**, alors que $\mathcal{H}_1$ est l'**hypothèse alternative**.

Exemple

Nous pouvons avoir comme hypothèse nulle : $\mathcal{H}_0 : \mu = \mu_0$ et, dans ce cas, une hypothèse alternative **pourrait être** $\mathcal{H}_1 : \mu \neq \mu_0$.

La décision d'un test consiste à choisir entre $\mathcal{H}_0$ et $\mathcal{H}_1$. Il y a donc quatre cas possibles qui sont détaillés dans le tableau ci-dessous :

	$\mathcal{H}_0$ vraie	$\mathcal{H}_1$ vraie
$\mathcal{H}_0$ décidée	$1 - \alpha$	β
$\mathcal{H}_1$ décidée	α	$1 - \beta$

où α et β sont les **risques d'erreur de première et de deuxième espèce**.

Définition 113.3

L'**erreur de première espèce** est le fait de rejeter l'hypothèse nulle $\mathcal{H}_0$ et donc de décider que l'hypothèse alternative $\mathcal{H}_1$ est vraie alors qu'en réalité, c'est l'hypothèse nulle $\mathcal{H}_0$ qui est vraie. Le **risque d'erreur** associé à cette décision est noté généralement α . Il s'agit donc de la probabilité de décider à tort que l'hypothèse alternative $\mathcal{H}_1$ est vraie.

Définition 113.4

L'**erreur de deuxième espèce** est le fait de ne pas rejeter l'hypothèse nulle $\mathcal{H}_0$ et donc de décider que l'hypothèse nulle $\mathcal{H}_0$ est vraie alors qu'en réalité, c'est l'hypothèse alternative $\mathcal{H}_1$ qui est vraie. Le **risque d'erreur** associé à cette décision est noté généralement β . Il s'agit donc de la probabilité de décider à tort que l'hypothèse nulle $\mathcal{H}_0$ est vraie.

**Point méthode**

Les valeurs les plus courantes pour α sont 10 %, 5 % ou 1%. Le fait d'imposer α faible conduit à une règle de décision plus stricte. En effet, dans ce cas, la décision consiste à abandonner $\mathcal{H}_0$ dans des cas rarissimes, et à conserver plus souvent, à tort, $\mathcal{H}_0$.



La situation idéale serait que ces deux erreurs soient nulles mais ce n'est pas possible. Pire encore, ces deux erreurs sont antagonistes. En effet, plus α est grande (respectivement petite), plus β est petite (respectivement grande). Depuis les travaux de Neyman et Pearson, l'erreur de première espèce est limitée à un niveau dit **niveau de significativité**.

Définition 113.5

La **puissance d'un test** est égale à $1 - \beta$ ou encore la puissance est la probabilité de rejeter $\mathcal{H}_0$ à raison. Elle représente la capacité du test à détecter $\mathcal{H}_1$.

**Point méthode**

Généralement la puissance doit au moins être égale à 0,80 pour être considérée comme satisfaisante.

2. Tests bilatéral et unilatéral

Avant d'appliquer un test statistique, il s'agit de bien définir le problème posé. En effet, selon les hypothèses formulées, nous appliquerons soit un **test bilatéral**, soit un **test unilatéral**.

Définition 113.6

Un **test bilatéral** s'applique quand nous cherchons une différence entre deux paramètres, ou entre un paramètre et une valeur donnée sans se préoccuper du signe ou du sens de la différence. Dans ce cas, la zone de rejet de l'hypothèse principale se fait de part et d'autre de la distribution de référence.

Définition 113.7

Un **test unilatéral** s'applique quand nous cherchons à savoir si un paramètre est supérieur (ou inférieur) à un autre ou à une valeur donnée. La zone de rejet de l'hypothèse principale est située d'un seul côté de la distribution de probabilité de référence.

Définition 113.8

Le risque d'erreur de première espèce α étant fixé, il faut choisir une **variable de décision** encore appelée **statistique de test**.



Cette variable est construite afin d'apporter de l'information sur le problème posé, à savoir le choix entre les deux hypothèses. Sa loi doit être parfaitement déterminée dans au moins une des deux hypothèses (le plus souvent dans $\mathcal{H}_0$) afin de ne pas introduire de nouvelles inconnues dans le problème.

Définition 113.9

La **région critique**, notée W (W pour wrong), encore appelée **zone de rejet** est égale à l'ensemble des valeurs de la variable de décision qui conduisent à écarter $\mathcal{H}_0$ au profit de $\mathcal{H}_1$. La **région critique** correspond donc aux intervalles dans lesquels les différences sont trop grandes pour être le fruit du hasard d'échantillonnage.



La région critique W peut être reliée au risque d'erreur de première espèce α par $\mathbb{P}_{\mathcal{H}_0}(W) \leq \alpha$, et lorsque la loi parente du modèle statistique est continue par $\mathbb{P}_{\mathcal{H}_0}(W) = \alpha$, où $\mathbb{P}_{\mathcal{H}_0}(W)$ est la probabilité que la statistique de test appartienne à W , sachant que $\mathcal{H}_0$ est vraie.

Définition 113.10

La **région d'acceptation**, notée $\overline{W}$, encore appelée **zone d'acceptation** est la région complémentaire de la région critique W . Elle correspond à l'intervalle dans lequel les différences observées entre les réalisations et la théorie sont attribuables aux fluctuations d'échantillonnage.



La région d'acceptation $\overline{W}$ peut être reliée au risque d'erreur de première espèce α par $\mathbb{P}_{\mathcal{H}_0}(\overline{W}) \geq 1 - \alpha$, et lorsque la loi parente du modèle statistique est continue par $\mathbb{P}_{\mathcal{H}_0}(\overline{W}) = 1 - \alpha$, où $\mathbb{P}_{\mathcal{H}_0}(\overline{W})$ est la probabilité que la statistique de test appartienne à $\overline{W}$, sachant que $\mathcal{H}_0$ est vraie.

Définition 113.11

La **p -valeur**, notée p et également appelée **valeur de p** , est la probabilité d'obtenir la même valeur, ou une valeur encore plus extrême dans la direction de l'hypothèse alternative $\mathcal{H}_1$, de la statistique de test si l'hypothèse nulle $\mathcal{H}_0$ était vraie.

3. La démarche à suivre pour la mise en place d'un test

► Comment réaliser un test et conclure à l'aide d'une région critique ?

1. Choix des deux hypothèses testées $\mathcal{H}_0$ et $\mathcal{H}_1$.
2. Détermination de la statistique du test.
3. Allure de la région critique en fonction de $\mathcal{H}_1$: test bilatéral ou unilatéral.
4. Calcul de la région critique en fonction de α .
5. Calcul de la valeur de la statistique de test observée sur l'échantillon.
6. Conclusion du test. Si la valeur calculée en 5 appartient à la région construite en 4, le test est significatif au niveau α . Nous rejetons l'hypothèse nulle $\mathcal{H}_0$ et nous décidons que l'hypothèse alternative $\mathcal{H}_1$ est vraie. Le risque d'erreur associé à cette décision est un risque d'erreur de première espèce qui vaut α . Si la valeur calculée en 5 n'appartient pas à la région construite en 4, le test n'est pas significatif au niveau α . Nous conservons l'hypothèse nulle $\mathcal{H}_0$ par défaut. Le risque d'erreur associé à cette décision est un risque d'erreur de deuxième espèce qui vaut β . Pour l'évaluer, il faudrait calculer la puissance $1 - \beta$ du test.
7. Calcul de la puissance $1 - \beta$ du test lorsque celui-ci n'est pas significatif.

► Comment réaliser un test et conclure à l'aide d'une p -valeur ?

1. Choix des deux hypothèses testées $\mathcal{H}_0$ et $\mathcal{H}_1$.
2. Détermination de la statistique du test.
3. Calcul de la p -valeur à partir des données de l'échantillon.

4. Conclusion du test. Si la p -valeur est inférieure ou égale à α , le test est significatif au niveau α . Nous rejetons l'hypothèse nulle $\mathcal{H}_0$ et nous décidons que l'hypothèse alternative $\mathcal{H}_1$ est vraie. Le risque d'erreur associé à cette décision est un risque d'erreur de première espèce qui vaut α . Si la p -valeur est strictement supérieure à α , le test n'est pas significatif au niveau α . Nous conservons l'hypothèse nulle $\mathcal{H}_0$ par défaut. Le risque d'erreur associé à cette décision est un risque d'erreur de deuxième espèce qui vaut β . Pour l'évaluer, il faudrait calculer la puissance $1 - \beta$ du test.
5. Calcul de la puissance $1 - \beta$ du test lorsque celui-ci n'est pas significatif.

4. Choix d'un test

Plusieurs tests de conception très différente sont souvent disponibles pour soumettre une hypothèse à une épreuve de vérité.

Définition 113.12

Le **test le plus puissant** est le test qui fournit, pour une même valeur de α , l'erreur β la plus petite ou encore la plus grande valeur de la puissance $1 - \beta$.



1. Le test le plus puissant détecte les plus petites différences entre les populations sans pour autant augmenter α .
2. Nous pouvons disposer de plusieurs tests pour vérifier une même hypothèse. En fonction du contexte, il faudra penser à utiliser le plus puissant d'entre eux.



Point méthode

- Les tests peu puissants augmentent la probabilité de commettre une erreur de deuxième espèce. Or, cette erreur peut s'avérer particulièrement grave.
- Pour évaluer la puissance d'un test, nous pourrions être amenés à utiliser des **courbes de puissance** parfois appelées **abaques**.

Définition 113.13

La **robustesse d'un test** équivaut à sa tolérance vis-à-vis du respect des conditions d'application du test.



La majorité des tests statistiques repose sur le respect d'un certain nombre de conditions. Selon le degré de respect de ces conditions d'utilisation, la validité des résultats se trouve plus ou moins affectée et elle l'est d'autant plus que le test est moins robuste.

Application

Considérons une analyse statistique en médecine qui permettrait de décider si un sujet est sain ($\mathcal{H}_0$) ou malade ($\mathcal{H}_1$).

1. Quel est le type d'erreur commise en classant comme malade un sujet bien portant ?
2. Quel est le type d'erreur commise en classant comme bien portant un sujet malade ?

-
1. C'est une erreur de première espèce.
 2. C'est une erreur de deuxième espèce.

Tests de conformité d'une espérance ou d'une variance à une norme

Rappelons que $\widehat{\mu}_n$ est défini à la Fiche 109, que $S_{n,c}^2$ est défini à la Fiche 110 et que $\widehat{\sigma}_n^2$ est défini à la Fiche 112.

1. Test d'une espérance d'une loi normale de variance connue

► Hypothèses testées

Nous souhaitons choisir entre les deux hypothèses suivantes :

$$\mathcal{H}_0 : \mu = \mu_0$$

contre

$$\mathcal{H}_1 : \mu \neq \mu_0.$$

► Conditions d'application du test

L'échantillon $x_1, \dots, x_n$ est formé de réalisations indépendantes de la variable aléatoire X qui suit une loi normale d'espérance μ inconnue et de variance σ^2 connue. Ce cas est peu fréquent.

► Statistique du test

Sous l'hypothèse nulle $\mathcal{H}_0$, la variable aléatoire $Z_n = \frac{\widehat{\mu}_n - \mu_0}{\sigma / \sqrt{n}}$ suit la loi normale $\mathcal{N}(0 ; 1)$.

 Fiche 112

► Règle de décision et conclusion du test

La valeur critique du test, notée $c_{1-\alpha/2}$, est lue dans une table de la loi normale centrée et réduite. Si la valeur absolue de la valeur de la statistique calculée sur l'échantillon, notée $Z_n(obs)$, est supérieure ou égale à $c_{1-\alpha/2}$, alors le test est significatif. Nous rejetons $\mathcal{H}_0$ et nous décidons que $\mathcal{H}_1$ est vraie avec un risque de première espèce α . Si la valeur absolue de la valeur de la statistique calculée sur l'échantillon, notée $Z_n(obs)$, est strictement inférieure à $c_{1-\alpha/2}$, alors le test n'est pas significatif. Nous conservons $\mathcal{H}_0$ avec un risque de deuxième espèce β à déterminer.

2. Test d'une espérance d'une loi normale de variance inconnue : le test t de Student

► Hypothèses testées

Nous souhaitons choisir entre les deux hypothèses suivantes :

$$\mathcal{H}_0 : \mu = \mu_0$$

contre

$$\mathcal{H}_1 : \mu \neq \mu_0.$$

► Conditions d'application du test

L'échantillon $x_1, \dots, x_n$ est formé de réalisations indépendantes de la variable aléatoire X qui suit une loi normale d'espérance μ et de variance σ^2 inconnues.

► Statistique du test

Sous l'hypothèse nulle $\mathcal{H}_0$, la variable aléatoire $T_{n-1} = \frac{\widehat{\mu}_n - \mu_0}{S_{n,c}/\sqrt{n}}$ suit la loi de Student $t(n-1)$, d'après le Théorème de Fisher.

 Fiche 112

► Règle de décision et conclusion du test

La valeur critique du test, notée $c_{1-\alpha/2}$, est lue dans une table de la loi de Student. Si la valeur absolue de la valeur de la statistique calculée sur l'échantillon, notée $T_{n-1}(obs)$, est supérieure ou égale à $c_{1-\alpha/2}$, alors le test est significatif. Nous rejetons $\mathcal{H}_0$ et nous décidons que $\mathcal{H}_1$ est vraie avec un risque de première espèce α . Si la valeur absolue de la valeur de la statistique calculée sur l'échantillon, notée $T_{n-1}(obs)$, est strictement inférieure à $c_{1-\alpha/2}$, alors le test n'est pas significatif. Nous conservons $\mathcal{H}_0$ avec un risque de deuxième espèce β à déterminer.

3. Test d'une variance de loi normale d'espérance connue

► Hypothèses testées

Nous souhaitons choisir entre les deux hypothèses suivantes :

$$\mathcal{H}_0 : \sigma^2 = \sigma_0^2$$

contre

$$\mathcal{H}_1 : \sigma^2 \neq \sigma_0^2.$$

► Conditions d'application du test

Il faut que l'échantillon $x_1, \dots, x_n$ soit des réalisations indépendantes de la variable aléatoire X qui suit une loi normale d'espérance μ connue et de variance σ^2 inconnue. Ce cas est peu fréquent.

► Statistique du test

Sous l'hypothèse nulle $\mathcal{H}_0$, la variable aléatoire $\chi_n^2 = \frac{n\widehat{\sigma}_n^2}{\sigma_0^2}$ suit la loi du Khi-deux $\chi^2(n)$.

 Fiche 112

► Règle de décision et conclusion du test

Les valeurs critiques du test, $c_{\alpha/2} < c_{1-\alpha/2}$, sont lues dans une table de la loi du Khi-deux. Si la valeur de la statistique calculée sur l'échantillon, notée $\chi_n^2(obs)$, n'appartient pas à l'intervalle $]c_{\alpha/2}; c_{1-\alpha/2}[$, le test est significatif. Nous rejetons l'hypothèse nulle $\mathcal{H}_0$ et nous décidons que l'hypothèse alternative $\mathcal{H}_1$ est vraie avec un risque de première espèce α . Si la valeur de la statistique calculée sur l'échantillon, notée $\chi_n^2(obs)$, appartient à l'intervalle $]c_{\alpha/2}; c_{1-\alpha/2}[$, le test n'est pas significatif. Nous conservons l'hypothèse nulle $\mathcal{H}_0$ avec un risque de deuxième espèce β à déterminer.

4. Test d'une variance de loi normale d'espérance inconnue

► Hypothèses testées

Nous souhaitons choisir entre les deux hypothèses suivantes :

$$\mathcal{H}_0 : \sigma^2 = \sigma_0^2$$

contre

$$\mathcal{H}_1 : \sigma^2 \neq \sigma_0^2.$$

► Conditions d'application du test

Il faut que l'échantillon $x_1, \dots, x_n$ soit des réalisations indépendantes de la variable aléatoire X qui suit une loi normale d'espérance μ et de variance σ^2 inconnues.

► Statistique du test

Sous l'hypothèse nulle $\mathcal{H}_0$, la v.a. $\chi_{n-1,c}^2 = \frac{(n-1)S_{n,c}^2}{\sigma_0^2}$ suit la loi du Khi-deux $\chi^2(n-1)$.



$$(n-1)S_{n,c}^2 = nS_c^2.$$

► Règle de décision et conclusion du test

Les valeurs critiques du test, $c_{\alpha/2} < c_{1-\alpha/2}$, sont lues dans une table de la loi du Khi-deux. Si la valeur de la statistique calculée sur l'échantillon, notée $\chi_{n-1,c}^2(obs)$, n'appartient pas à l'intervalle $]c_{\alpha/2}; c_{1-\alpha/2}[$, alors le test est significatif. Nous rejetons l'hypothèse nulle $\mathcal{H}_0$ et nous décidons que l'hypothèse alternative $\mathcal{H}_1$ est vraie avec un risque de première espèce α . Si la valeur de la statistique calculée sur l'échantillon, notée $\chi_{n-1,c}^2(obs)$, appartient à l'intervalle $]c_{\alpha/2}; c_{1-\alpha/2}[$, alors le test n'est pas significatif. Nous conservons l'hypothèse nulle $\mathcal{H}_0$ avec un risque de deuxième espèce β à déterminer.

Application Gaz nocif

Dans l'atmosphère, le taux d'un gaz nocif, pour un volume donné, suit une loi normale d'espérance μ et de variance σ^2 . Nous effectuons n prélèvements conduisant aux valeurs $x_1, \dots, x_n$.

1. Nous savons que $\sigma^2 = 100$ et que sur $n = 10$ prélèvements, nous avons trouvé une valeur moyenne de 52. Pouvons-nous conclure au risque $\alpha = 5 \%$ que l'espérance μ est inférieure à 50, qui est le seuil tolérable admis ? Pouvons-nous donner cette conclusion au risque $\alpha = 10 \%$?

2. Nous ne connaissons pas la variance σ^2 , mais nous avons effectué 200 prélèvements. Nous trouvons une valeur moyenne égale à 51 et une variance empirique corrigée égale à 100. Pouvons-nous conclure au risque $\alpha = 5 \%$ que l'espérance μ est inférieure à 50, qui est le seuil tolérable admis ?

1. Compte tenu du problème posé nous optons pour un test **unilatéral à droite** :

$$\mathcal{H}_0 : \mu = \mu_0 \quad (\leq) \quad \text{contre} \quad \mathcal{H}_1 : \mu > \mu_0.$$

L'écart-type de la population, $\sigma = 10$, est connu : nous utilisons la loi normale. La zone de rejet est $[1,645; +\infty[$ puisque $\mathbb{P}[Z \geq 1,645] = 1 - 0,95 = 0,05$, où Z suit la loi normale centrée et réduite. La valeur 1,645 est lue dans le tableau fourni Fiche 121. La réalisation de la statistique du test est $\sqrt{10} \times \frac{52 - 50}{10} = \frac{\sqrt{10}}{5} \approx 0,63 < 1,645$. Le test n'est donc pas significatif au risque $\alpha = 5 \%$. Nous décidons, qu'au risque $\alpha = 5 \%$, l'espérance μ est inférieure ou égale à 50. Le type d'erreur associée à cette décision est une erreur de deuxième espèce qu'il faudrait évaluer avant de pouvoir conclure que notre décision est correcte.

Au risque $\alpha = 10 \%$, la zone de rejet devient $[1,282; +\infty[$ puisque $\mathbb{P}[Z \geq 1,282] = 0,10$. La valeur 1,282 est lue dans le tableau fourni Fiche 121. Comme $0,63 < 1,282$, le test n'est donc pas significatif au seuil $\alpha = 10 \%$. Mêmes conclusions qu'en 1..

2. Compte tenu du problème posé nous optons pour un test **unilatéral à droite** :

$$\mathcal{H}_0 : \mu = \mu_0 \quad (\leq) \quad \text{contre} \quad \mathcal{H}_1 : \mu > \mu_0.$$

L'écart-type σ de la population est inconnu : nous utilisons la loi de Student à $200 - 1 = 199$ degrés de liberté. La zone de rejet est $[1,6525; +\infty[$ puisque $t_{199;0,95} = 1,6525$, commande `qt(0.95,199)` avec le langage **R**. La réalisation de la statistique du test est $\sqrt{200} \times \frac{51 - 50}{10} = \sqrt{2} \approx 1,4142 < 1,6525$. Le test n'est donc pas significatif au risque $\alpha = 5 \%$. Nous concluons, avec un risque d'erreur de première espèce de 5% , que l'espérance μ est inférieure ou égale à 50.

Fiche 121

Fiche 121

Pour s'entraîner (solutions p. 554)

114.1 Un échantillon de taille $n = 100$ a été prélevé au hasard dans une population normale de moyenne μ et d'écart-type σ . La moyenne et la variance corrigée calculées sur cet échantillon sont $\widehat{\mu}_{100}(obs) = 4$ et $S_{100,c}^2(obs) = 9$.

1. Tester l'hypothèse $\sigma^2 = 8$ au seuil $\alpha = 5 \%$. Pour ce test, vous utiliserez $c_{0,025} = 73,36$ et $c_{0,975} = 128,42$.

2. En admettant que la variance σ^2 est connue et est égale à 8, tester l'hypothèse $\mu = 3$ au seuil $\alpha = 5 \%$.

3. Tester, au seuil $\alpha = 5 \%$, l'hypothèse $\mu = 3$ sans faire aucune hypothèse sur la valeur de σ^2 .

4. En admettant et en utilisant le fait que l'espérance μ est connue et est égale à 3, tester, au seuil $\alpha = 5 \%$, l'hypothèse $\sigma^2 = 8$? Pour ce test, vous utiliserez $c_{0,025} = 74,22$ et $c_{0,975} = 129,56$.

114.2 Les mesures de la vitesse de la lumière réalisées par Michelson et Morley, reproduites à la Fiche 109, sont-elles compatibles avec la valeur de 299734,5 km/s qui est la vitesse de la lumière dans l'air communément acceptée de nos jours au seuil $\alpha = 5 \%$? Pour ce test, vous utiliserez $c_{0,975} = 1,9842$.

Tests de comparaison de deux variances

1. Comparons deux variances : deux échantillons indépendants

► Lois normales d'espérances connues

Ce cas est peu fréquent dans la pratique. C'est pourquoi il ne sera pas exposé dans cet ouvrage. Pour l'utiliser, il faudra donc adapter le cas ci-dessous au cas où les espérances sont inconnues.

► Lois normales d'espérances inconnues : le test de Fisher-Snedecor

Hypothèses testées

Nous souhaitons choisir entre les deux hypothèses suivantes :

$$\mathcal{H}_0 : \sigma_1^2 = \sigma_2^2$$

contre

$$\mathcal{H}_1 : \sigma_1^2 \neq \sigma_2^2.$$

Conditions d'application du test

Il faut que l'échantillon $x_1, \dots, x_{n_1}$ soit des réalisations indépendantes de la variable aléatoire X qui suit la loi normale $\mathcal{N}(\mu_1; \sigma_1)$ et que le second échantillon $y_1, \dots, y_{n_2}$ soit des réalisations indépendantes de la variable aléatoire Y qui suit la loi normale $\mathcal{N}(\mu_2; \sigma_2)$. De plus n_1 et n_2 ne sont pas forcément égaux et les variables aléatoires X et Y doivent être indépendantes.

Statistique du test

Déterminer si $\sigma_1^2 = \sigma_2^2$ est équivalent à déterminer si $\sigma_1^2/\sigma_2^2 = 1$, ce qui explique la forme de la statistique utilisée pour le test. Sous l'hypothèse nulle $\mathcal{H}_0$, la variable aléatoire $F = S_{n_1,c}^2/S_{n_2,c}^2$, où $S_{n,c}^2$ est défini à la Fiche 110, suit la loi de Fisher $F(n_1 - 1, n_2 - 1)$.

 Fiche 110

Règle de décision et conclusion du test

Les valeurs critiques du test, $c_{\alpha/2}$ et $c_{1-\alpha/2}$, sont lues dans une table de la loi de Fisher. Si la valeur de la statistique calculée sur l'échantillon, notée $F(obs)$, n'appartient pas à l'intervalle $]c_{\alpha/2}; c_{1-\alpha/2}[$, alors le test est significatif. Nous rejetons $\mathcal{H}_0$ et nous décidons que $\mathcal{H}_1$ est vraie avec un risque de première espèce α . Si la valeur de la statistique calculée sur l'échantillon, notée $F(obs)$, appartient à l'intervalle $]c_{\alpha/2}; c_{1-\alpha/2}[$, alors le test n'est pas significatif. Nous conservons $\mathcal{H}_0$ avec un risque de deuxième espèce β à déterminer.



Les décisions des tests unilatéraux se déduisent facilement de la procédure que nous venons d'introduire.

2. Comparons deux variances : deux échantillons dépendants

► Lois normales d'espérances connues

Ce cas est peu fréquent dans la pratique. C'est pourquoi il ne sera pas exposé dans cet ouvrage. Pour l'utiliser, il faudra donc adapter le cas ci-dessous au cas où les espérances sont inconnues.

► Lois normales d'espérances inconnues

Hypothèses testées

Nous souhaitons choisir entre les deux hypothèses suivantes :

$$\mathcal{H}_0 : \sigma_1^2 = \sigma_2^2$$

contre

$$\mathcal{H}_1 : \sigma_1^2 \neq \sigma_2^2.$$

Conditions d'application du test

Il faut que l'échantillon $(x_1; y_1), \dots, (x_n; y_n)$ soit des réalisations indépendantes du couple de variables aléatoires (X, Y) qui suit une loi normale à deux dimensions. La loi de X est une loi $\mathcal{N}(\mu_1; \sigma_1)$ et celle de Y est une loi $\mathcal{N}(\mu_2; \sigma_2)$. Les deux échantillons $x_1, \dots, x_n$ et $y_1, \dots, y_n$ sont donc appariés et leurs effectifs sont nécessairement égaux.

Principe du test

Le test est basé sur la relation entre la différence des variances des variables aléatoires X et Y et la covariance des variables aléatoires $X + Y$ et $X - Y$: $\mathbb{V}\text{ar}(X) - \mathbb{V}\text{ar}(Y) = \mathbb{C}\text{ov}(X + Y, X - Y)$.

Ainsi tester l'égalité des variances des variables aléatoires X et Y , $\mathbb{V}\text{ar}(X) - \mathbb{V}\text{ar}(Y) = 0$, est équivalent à tester si la corrélation linéaire entre $X + Y$ et $X - Y$ est nulle :

$$\rho(X + Y, X - Y) = \frac{\mathbb{C}\text{ov}(X + Y, X - Y)}{\sigma(X + Y) \times \sigma(X - Y)}.$$

Statistique du test

Nous utilisons l'estimateur $\rho(\widehat{X}, \widehat{Y})$ du coefficient de corrélation linéaire $\rho(X, Y)$ défini par :

$$\rho(\widehat{X}, \widehat{Y}) = \frac{\sum_{i=1}^n (X_i - \widehat{\mu}_n(X))(Y_i - \widehat{\mu}_n(Y))}{(n-1) \sqrt{S_{n,c}^2(X) S_{n,c}^2(Y)}} = \frac{\sum_{i=1}^n X_i Y_i - \widehat{\mu}_n(X) \widehat{\mu}_n(Y)}{\sqrt{\sum_{i=1}^n X_i^2 - \widehat{\mu}_n(X)^2} \sqrt{\sum_{i=1}^n Y_i^2 - \widehat{\mu}_n(Y)^2}}.$$

Si l'hypothèse nulle $\mathcal{H}_0$ est vraie et lorsque les conditions d'utilisation sont vérifiées, la

variable aléatoire $\mathcal{R}_{n-2} = \sqrt{n-2} \frac{\rho(\widehat{X} + \widehat{Y}, \widehat{X} - \widehat{Y})}{\sqrt{1 - \rho(\widehat{X} + \widehat{Y}, \widehat{X} - \widehat{Y})^2}}$ suit la loi de Student $t(n-2)$.

Règle de décision et conclusion du test

La valeur critique du test, notée $c_{1-\alpha/2}$, est lue dans une table de la loi de Student. Si la valeur absolue de la valeur de la statistique calculée sur l'échantillon, notée $\mathcal{R}_{n-2}(obs)$, est supérieure ou égale à $c_{1-\alpha/2}$, alors le test est significatif. Nous rejetons $\mathcal{H}_0$ et nous décidons que $\mathcal{H}_1$ est vraie avec un risque de première espèce α . Si la valeur absolue de la valeur de la statistique calculée sur l'échantillon, notée $\mathcal{R}_{n-2}(obs)$, est strictement inférieure à $c_{1-\alpha/2}$, alors le test n'est pas significatif. Nous conservons $\mathcal{H}_0$ avec un risque de deuxième espèce β à déterminer.



Les décisions des tests unilatéraux se déduisent facilement de la procédure que nous venons d'introduire.

Application

1. Le Coucou gris adapte-t-il ses œufs en fonction de la taille du nid ?

Le Coucou gris¹ est un oiseau qui fait couvrir ses œufs par des oiseaux d'autres espèces que la sienne, de tailles très différentes. Une hypothèse a été émise selon laquelle le Coucou pourrait adapter la taille de ses œufs à la taille du nid dans lequel il pond. Une étude faite par le biologiste O.H. Latter et publiée sous le titre « The Egg Of Cuculus Canorus : An Enquiry Into The Dimensions Of The Cuckoo'S Egg And The Relation Of The Variations To The Size Of The Eggs Of The Foster-Parent, With Notes On Coloration. » dans la revue *Biometrika*, 1902, Volume 1, pages 164-176, sur la taille des œufs déposés dans les nids de petite taille (Roitelet) ou de plus grande taille (Fauvette) a donné les valeurs (en millimètres) suivantes :

1. Dans les nids de Roitelet : 19, 8 ; 22, 1 ; 21, 5 ; 20, 9 ; 22, 0 ; 21, 0 ; 22, 3 ; 21, 0 ; 20, 3 ; 20, 9 ; 22, 0 ; 22, 0 ; 20, 8 ; 21, 2 ; 21, 0.
2. Dans les nids de Fauvette : 22, 0 ; 23, 9 ; 20, 9 ; 23, 8 ; 25, 0 ; 24, 0 ; 23, 8 ; 21, 7 ; 22, 8 ; 23, 1 ; 23, 5 ; 23, 0 ; 23, 0 ; 23, 1.

Au risque $\alpha = 5\%$, les deux populations ont-elles la même variance ?

2. Mesure de résistance.

Pour vérifier si deux ohmmètres Ω_X et Ω_Y permettent d'obtenir des résultats ayant une précision identique, les mêmes onze résistances ont été mesurées à l'aide des deux appareils. Des expériences antérieures ont permis de montrer que l'échantillon $((x_1, y_1), \dots, (x_{11}, y_{11}))$ est formé de réalisations d'une loi normale à deux dimensions. Que conclure à partir des résultats suivants ?

Ω_X	99,48	98,48	101,01	102,07	98,18	100,08	101,15	97,76	101,39	95,00	99,06
Ω_Y	98,59	97,91	100,51	100,73	97,56	99,09	101,17	98,72	100,43	97,88	99,57

Pour répondre à la question posée, nous cherchons à tester les deux hypothèses suivantes :

$$\mathcal{H}_0 : \sigma_1^2 = \sigma_2^2$$

contre

$$\mathcal{H}_1 : \sigma_1^2 \neq \sigma_2^2.$$



Nous allons utiliser le test de Fisher-Snedecor de comparaison des variances. Or ce dernier repose sur une hypothèse de normalité, que nous devons commencer par vérifier.

Pour la variable « taille des œufs de Coucou dans les nids de Roitelet », la statistique W du test de Shapiro-Wilk vaut 0,9295 pour une valeur critique $w_{0,05} = 0,881$ et, pour la variable



Fiche 118



Fiche 121

« taille des œufs de Coucou dans les nids de Fauvette », la statistique W du test de Shapiro-Wilk vaut 0,9558 pour une valeur critique $w_{0,05} = 0,874$. Aucun des deux tests n'est donc significatif au risque $\alpha = 5\%$. Nous conservons l'hypothèse nulle $\mathcal{H}_0$ dans les deux cas.



Le risque d'erreur associé à chacune de ces décisions est un risque d'erreur de deuxième espèce β qui est considéré comme raisonnable dans ce cas particulier, malgré l'effectif réduit des deux échantillons, en raison des autres mesures présentées dans l'article de O.H. Latter.

La variance corrigée de l'échantillon des Roitelets est $S_{15,c}^2(obs) = 0,5155$; celle de l'échantillon des Fauvettes est $S_{14,c}^2(obs) = 1,1013$. Les valeurs critiques du test de Fisher-Snedecor de comparaison des variances sont celles d'une loi de Fisher-Snedecor à 14 degrés de liberté au numérateur et 13 degrés de liberté au dénominateur : $c_{0,025;14;13} = 1/c_{0,975;13;14} = 1/3,0119 = 0,3320$ et $c_{0,975;14;13} = 3,0819$. Ainsi, la réalisation de la statistique du test $F(obs) = \frac{0,5155}{1,1013} = 0,4681$ appartient à $]0,3320; 3,0819[$ et le test n'est pas significatif au risque $\alpha = 5\%$. Nous conservons par défaut l'hypothèse nulle $\mathcal{H}_0 : \sigma_1^2 = \sigma_2^2$, écrite ci-dessus. Le risque d'erreur associé à cette décision est un risque de deuxième espèce β à déterminer.

2. Nous obtenons $\rho(X + \widehat{Y}, \bar{X} - Y)(obs) = 0,6608$ et $\sqrt{9} \times 0,6608 / \sqrt{1 - 0,6608^2} = 2,6411 > 2,2622$ où 2,2622 est le quantile d'ordre 0,975 de la loi de Student à 9 degrés de liberté $t(9)$. Le test est significatif au risque $\alpha = 5\%$. Il y a une différence entre les précisions des deux ohmmètres.

Fiche 121

Fiche 121

Fiche 121

Pour s'entraîner (solutions p. 555)

115.1 Nous avons prélevé au hasard un échantillon de taille $n_1 = 10$ dans une population normale de moyenne μ_1 et d'écart-type σ_1 . La moyenne et la variance corrigée calculées sur cet échantillon sont respectivement égales à $\widehat{\mu}_{n_1}(obs) = 4$ et $S_{n_1,c}^2(obs) = 6$. Nous avons aussi prélevé au hasard un second échantillon de taille $n_2 = 15$ dans une population normale de moyenne μ_2 et d'écart-type σ_2 . La moyenne et la variance corrigée calculées sur cet échantillon sont respectivement égales à $\widehat{\mu}_{n_2}(obs) = 7$ et $S_{n_2,c}^2(obs) = 20$. Tester l'hypothèse $\sigma_1^2 = \sigma_2^2$ au risque $\alpha = 5\%$.

115.2

www

Il y a des raisons de penser que l'épaisseur de la cire dont sont enduits des sacs en papier est plus irrégulière à l'intérieur qu'à l'extérieur. Pour le vérifier 75 mesures de l'épaisseur ont été faites et ont donné les résultats suivants :

- surface intérieure : $\sum_{i=1}^{75} x_i = 72$ et $\sum_{i=1}^{75} x_i^2 = 91$.
- surface extérieure : $\sum_{i=1}^{75} y_i = 49$ et $\sum_{i=1}^{75} y_i^2 = 84$.
- $\sum_{i=1}^{75} x_i y_i = 38$.

Faire un test pour déterminer, au risque $\alpha = 5\%$, si la variabilité de l'épaisseur de la cire est plus grande à l'intérieur qu'à l'extérieur des sacs.

Tests de comparaison de deux espérances

Rappelons que $\widehat{\mu}_n$ est défini Fiche 109 et que S_n^2 est défini Fiche 110.

1. Comparons deux espérances de lois normales de variances connues

Soient X une variable aléatoire qui suit la loi normale $\mathcal{N}(\mu_1; \sigma_1)$ et Y une variable aléatoire qui suit la loi normale $\mathcal{N}(\mu_2; \sigma_2)$ avec σ_1 et σ_2 connus.

► Hypothèses testées

Nous souhaitons choisir entre les deux hypothèses suivantes :

$$\mathcal{H}_0 : \mu_1 = \mu_2$$

contre

$$\mathcal{H}_1 : \mu_1 \neq \mu_2.$$

► Conditions d'application du test

Il faut que l'échantillon $x_1, \dots, x_{n_1}$ soit des réalisations indépendantes de la variable aléatoire X qui suit une loi normale, que le second échantillon $y_1, \dots, y_{n_2}$ soit aussi des réalisations indépendantes de la variable aléatoire Y qui suit une loi normale et que les variables aléatoires X et Y soient indépendantes. De plus, les effectifs n_1 et n_2 peuvent ne pas être égaux.

 Fiche 103

► Statistique du test

Sous l'hypothèse nulle $\mathcal{H}_0$, la variable aléatoire $Z_{n_1, n_2} = \frac{\widehat{\mu}_{n_1} - \widehat{\mu}_{n_2}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$ suit la loi normale $\mathcal{N}(0; 1)$.

► Règle de décision et conclusion du test

La valeur critique du test, notée $c_{1-\alpha/2}$, est lue dans une table de la loi normale centrée et réduite. Si la valeur absolue de la valeur de la statistique calculée sur l'échantillon, notée $Z_{n_1, n_2}(obs)$, est supérieure ou égale à $c_{1-\alpha/2}$, alors le test est significatif. Nous rejetons $\mathcal{H}_0$ et nous décidons que $\mathcal{H}_1$ est vraie avec un risque de première espèce α . Si la valeur absolue de la valeur de la statistique calculée sur l'échantillon, notée $Z_{n_1, n_2}(obs)$, est strictement inférieure à $c_{1-\alpha/2}$, alors le test n'est pas significatif. Nous conservons $\mathcal{H}_0$ avec un risque de deuxième espèce β à déterminer.

2. Comparons deux espérances de lois normales de variances inconnues et égales : le test t de Student

Soient X une variable aléatoire qui suit la loi normale $\mathcal{N}(\mu_1; \sigma_1)$ et Y une variable aléatoire qui suit la loi normale $\mathcal{N}(\mu_2; \sigma_2)$ avec σ_1 et σ_2 inconnus.

► Hypothèses testées

Ce sont les mêmes hypothèses que dans la première section.

► Conditions d'application du test

Il faut que l'échantillon $x_1, \dots, x_{n_1}$ soit des réalisations indépendantes de la variable aléatoire X qui suit une loi normale, que le second échantillon $y_1, \dots, y_{n_2}$ soit aussi des réalisations indépendantes de la variable aléatoire Y qui suit une loi normale et que les variables aléatoires X et Y soient indépendantes. De plus, les effectifs n_1 et n_2 peuvent ne pas être égaux. Enfin, il faut que

$$\sigma_1^2 = \sigma_2^2 = \sigma^2.$$

Une question se pose alors naturellement : « Comment vérifier que cette hypothèse est justifiée ? ». Ce sera le cas, si le résultat de l'application du test de Fisher-Snedecor aux résultats expérimentaux observés aboutit à conserver l'hypothèse nulle $\mathcal{H}_0$ d'égalité des variances.

 Fiche 115

► Statistique du test

Sous l'hypothèse nulle $\mathcal{H}_0$, la variable aléatoire $T_{n_1+n_2-2} = \frac{\widehat{\mu}_{n_1} - \widehat{\mu}_{n_2}}{\widehat{\sigma} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$, où $\widehat{\sigma} = \sqrt{\frac{n_1 S_{n_1}^2 + n_2 S_{n_2}^2}{n_1 + n_2 - 2}}$, suit la loi de Student $t(n_1 + n_2 - 2)$.

► Règle de décision et conclusion du test

La valeur critique du test, notée $c_{1-\alpha/2}$, est lue dans une table de la loi de Student. Si la valeur absolue de la valeur de la statistique calculée sur l'échantillon, notée $T_{n_1+n_2-2}(obs)$, est supérieure ou égale à $c_{1-\alpha/2}$, alors le test est significatif. Nous rejetons $\mathcal{H}_0$ et nous décidons que $\mathcal{H}_1$ est vraie avec un risque de première espèce α . Si la valeur absolue de la valeur de la statistique calculée sur l'échantillon, notée $T_{n_1+n_2-2}(obs)$, est strictement inférieure à $c_{1-\alpha/2}$, alors le test n'est pas significatif. Nous conservons $\mathcal{H}_0$ avec un risque de deuxième espèce β à déterminer.

3. Comparons deux espérances de lois normales de variances inconnues quelconques : le test de Welch

Soit X une variable aléatoire qui suit la loi normale $\mathcal{N}(\mu_1; \sigma_1)$ et Y une variable aléatoire qui suit la loi normale $\mathcal{N}(\mu_2; \sigma_2)$ avec σ_1 et σ_2 inconnus.

► Hypothèses testées

Ce sont les mêmes hypothèses que dans la première section.

► Conditions d'application du test

Il faut que l'échantillon $x_1, \dots, x_{n_1}$ soit des réalisations indépendantes de la variable aléatoire X qui suit une loi normale, que le second échantillon $y_1, \dots, y_{n_2}$ soit aussi des réalisations indépendantes de la variable aléatoire Y qui suit une loi normale et les variables aléatoires X et Y soient indépendantes. De plus, les effectifs n_1 et n_2 , de même

que les variances σ_1^2 et σ_2^2 , peuvent ne pas être égaux. Ainsi, si le test de Fisher-Snedecor aboutit au rejet de l'hypothèse nulle $\mathcal{H}_0$, il est nécessaire d'utiliser ce test de Welch à la place du test t de Student.

► Statistique du test

Sous l'hypothèse nulle $\mathcal{H}_0$, la variable aléatoire $T_\nu = \frac{\widehat{\mu}_{n_1} - \widehat{\mu}_{n_2}}{\sqrt{\frac{S_{n_1}^2}{n_1 - 1} + \frac{S_{n_2}^2}{n_2 - 1}}}$ suit la loi de Student $t(\nu)$, où ν est l'entier le plus proche de $\left(\frac{S_{n_1}^2}{n_1 - 1} + \frac{S_{n_2}^2}{n_2 - 1} \right)^2 \left/ \left(\frac{S_{n_1}^4}{(n_1 - 1)^3} + \frac{S_{n_2}^4}{(n_2 - 1)^3} \right) \right.$.

► Règle de décision et conclusion du test

La valeur critique du test, notée $c_{1-\alpha/2}$, est lue dans une table de la loi de Student. Si la valeur absolue de la valeur de la statistique calculée sur l'échantillon, notée $T_\nu(obs)$, est supérieure ou égale à $c_{1-\alpha/2}$, alors le test est significatif. Nous rejetons $\mathcal{H}_0$ et nous décidons que $\mathcal{H}_1$ est vraie avec un risque de première espèce α . Si la valeur absolue de la valeur de la statistique calculée sur l'échantillon, notée $T_\nu(obs)$, est strictement inférieure à $c_{1-\alpha/2}$, alors le test n'est pas significatif. Nous conservons $\mathcal{H}_0$ avec un risque de deuxième espèce β à déterminer.

4. Comparons deux espérances de lois quelconques de variances inconnues

Soit X une variable aléatoire d'espérance égale à μ_1 et d'écart-type égal à σ_1 et Y une variable aléatoire d'espérance égale à μ_2 et d'écart-type égal à σ_2 avec σ_1 et σ_2 inconnus.

► Hypothèses testées

Ce sont les mêmes hypothèses que dans la première section.

► Conditions d'application du test

Les effectifs n_1 et n_2 sont tous les deux supérieurs à 30.

► Statistique du test

Sous l'hypothèse nulle $\mathcal{H}_0$ et si les conditions d'application du test sont remplies, la variable aléatoire $\zeta_{n_1, n_2} = \frac{\widehat{\mu}_{n_1} - \widehat{\mu}_{n_2}}{\sqrt{\frac{S_{n_1}^2}{n_1 - 1} + \frac{S_{n_2}^2}{n_2 - 1}}}$ suit **approximativement** la loi normale

$\mathcal{N}(0; 1)$.

► Règle de décision et conclusion du test

La valeur critique du test, notée $c_{1-\alpha/2}$, est lue dans une table de la loi normale centrée et réduite. Si la valeur absolue de la valeur de la statistique calculée sur l'échantillon, notée $\zeta_{n_1, n_2}(obs)$, est supérieure ou égale à $c_{1-\alpha/2}$, alors le test est significatif. Nous rejetons $\mathcal{H}_0$ et nous décidons que $\mathcal{H}_1$ est vraie avec un risque de première espèce α . Si la valeur absolue de la valeur de la statistique calculée sur l'échantillon, notée $\zeta_{n_1, n_2}(obs)$, est

strictement inférieure à $c_{1-\alpha/2}$, alors le test n'est pas significatif. Nous conservons $\mathcal{H}_0$ avec un risque de deuxième espèce β à déterminer.

Application La taille du nid de coucou

Reprendre les données correspondantes de la Fiche 115. Déterminer, au seuil de significativité $\alpha = 5\%$, si le Coucou gris adapte la taille de ses œufs à la taille du nid dans lequel il les pond.



Nous cherchons à comparer la taille moyenne des œufs de coucou pondus dans les nids de Roitelets et à celle de ceux pondus dans les nids de Fauvettes. À la Fiche 115 nous avons montré, qu'au seuil de 5 %, nous ne pouvons rejeter l'hypothèse que les valeurs de chacun des échantillons sont des réalisations de variables aléatoires qui suivent des lois normales et que celles-ci ont la même variance.

La moyenne de l'échantillon des Roitelets est 21,2533 et celle de l'échantillon des Fauvettes est 23,1143. La variance non corrigée de l'échantillon des Roitelets est 0,4812. La variance non corrigée de l'échantillon des Fauvettes est 1,0227. La valeur critique du test de Student de comparaison des moyennes est celle d'une loi de Student à $15 + 14 - 2 = 27$ degrés de liberté : $c_{0,975} = 2,0518$.

Remarquons que $|T_{27}(obs)| = \frac{|21,2533 - 23,1143|}{\sqrt{\frac{15 \times 0,4812 + 14 \times 1,0227}{15 + 14 - 2}}} \sqrt{\frac{1}{15} + \frac{1}{14}} = 5,6074 \geq 2,0518$.

Par conséquent, le test est significatif au seuil $\alpha = 5\%$. Nous rejetons l'hypothèse nulle $\mathcal{H}_0$ et nous décidons que l'hypothèse alternative $\mathcal{H}_1$ est vraie avec un risque de première espèce $\alpha = 5\%$.



Nous pouvons donc affirmer que le Coucou gris adapte la grosseur de ses œufs en fonction de la taille du nid dans lequel il pond. Il semble qu'il s'agisse d'un phénomène de mimétisme qui permet aux œufs du Coucou gris de passer plus facilement inaperçus dans le nid dans lequel ils se trouvent.

Fiche 121

Pour s'entraîner (solutions p. 555)

116.1 Deux chaînes de fabrication produisent des transistors. Des relevés effectués pendant 10 jours ont donné les résultats suivants :

- ligne 1 : $n_1 = 10$, $\widehat{\mu}_{n_1}(obs) = 2800$ et $\sum_{i=1}^{n_1} (x_i - \widehat{\mu}_{n_1}(obs))^2 = 207200$;
- ligne 2 : $n_2 = 10$, $\widehat{\mu}_{n_2}(obs) = 2680$ et $\sum_{i=1}^{n_2} (y_i - \widehat{\mu}_{n_2}(obs))^2 = 38200$.

Nous savons que les deux distributions, dont sont issus les deux échantillons, sont normales.

Pouvons-nous conclure, au risque $\alpha = 5\%$, à une différence entre les deux chaînes de fabrication ?

Tests du Khi-deux : adéquation et indépendance

1. Test du Khi-deux : adéquation

► Contexte du test

Le test suivant permet de s'intéresser à la possibilité de l'adéquation de la loi d'une variable aléatoire X à une loi de probabilité donnée. Il est adapté pour des lois de probabilité discrètes et peut être également utilisé pour une loi continue **entièrement spécifiée**. Nous détaillons son application au cas de la loi discrète finie définie par $(q_{x_i})_{x_i \in \mathcal{X}}$ avec $\text{Card}(\mathcal{X}) = I$, $\forall x_i \in \mathcal{X}$, $q_{x_i} \in [0; 1]$ et $\sum_{i=1}^I q_{x_i} = 1$.

Hypothèses testées

$$\mathcal{H}_0 : (\mathbb{P}[X = x_i] = p_{x_i})_{x_i \in \mathcal{X}} = (q_{x_i})_{x_i \in \mathcal{X}}$$

contre

$$\mathcal{H}_1 : (\mathbb{P}[X = x_i] = p_{x_i})_{x_i \in \mathcal{X}} \neq (q_{x_i})_{x_i \in \mathcal{X}}.$$

► Procédure du test

Conditions d'application du test

Considérons un échantillon formé de réalisations indépendantes de la variable aléatoire X et de taille $n = m_{\bullet} = m_1 + \dots + m_I$ où m_i est l'effectif observé associé à x_i , c'est-à-dire le nombre d'individus de l'échantillon qui présentent x_i . Soit $c_i = m_{\bullet} q_{x_i}$ l'effectif théorique associé à x_i , c'est-à-dire le nombre d'individus que, compte tenu de la taille de l'échantillon et de la loi $(q_{x_i})_{x_i \in \mathcal{X}}$, nous nous attendions à voir présenter x_i . Des conditions portent sur les **effectifs théoriques**, nous devons avoir $c_i \geq 5$, et sur l'**effectif total de l'échantillon**, nous devons avoir $m_{\bullet} \geq 50$.

Statistique du test

Si l'hypothèse nulle $\mathcal{H}_0$ est vérifiée et lorsque les conditions d'application du test sont remplies, $\chi^2(\text{obs}) = \sum_{i=1}^I \frac{(m_i - c_i)^2}{c_i}$ est une réalisation d'une variable aléatoire qui suit approximativement la loi du Khi-deux à $(I - 1)$ degrés de liberté.

Règle de décision et conclusion du test

La valeur critique du test, notée $c_{1-\alpha}$, est lue dans une table de la loi du Khi-deux. Si la valeur de la statistique calculée sur l'échantillon, notée $\chi^2(\text{obs})$, est supérieure ou égale à $c_{1-\alpha}$, le test est significatif. Nous rejetons l'hypothèse nulle $\mathcal{H}_0$ et nous décidons que l'hypothèse alternative $\mathcal{H}_1$ est vraie avec un risque de première espèce α . Si la valeur de la statistique calculée sur l'échantillon, notée $\chi^2(\text{obs})$, est inférieure à $c_{1-\alpha}$, le test n'est pas significatif. Nous conservons l'hypothèse nulle $\mathcal{H}_0$ avec un risque de deuxième espèce β à déterminer.



Il est possible d'utiliser ce test pour tester l'adéquation possible à une loi appartenant à une famille de lois dont il faut de surcroît estimer les paramètres, au nombre de m . La difficulté principale de cette extension est qu'il **faut avoir conscience** que la technique d'estimation utilisée pour obtenir le vecteur m des paramètres estimés a une influence sur la nature de la loi asymptotique de la statistique du test. Si I classes sont utilisées et m paramètres sont estimés avec les estimateurs de maximum de vraisemblance pour la loi multinomiale, c'est-à-dire à partir du regroupement en classes, alors cette loi est la loi du Khi-deux à $I - m - 1$ degrés de liberté.

Fiches 86
et 108

2. Test du Khi-deux : indépendance

► Contexte du test

Le test du χ^2 d'indépendance sert à étudier la liaison entre deux caractères qualitatifs X et Y . Nous considérons donc le tableau suivant :

$X \backslash Y$	Modalité 1	...	Modalité j	...	Modalité J	Totaux
Modalité 1	$m_{1,1}$	...	$m_{1,j}$	...	$m_{1,J}$	$m_{1,\bullet}$
...	...	...	...	...	...	...
Modalité i	$m_{i,1}$	...	$m_{i,j}$	...	$m_{i,J}$	$m_{i,\bullet}$
...	...	...	...	...	...	...
Modalité I	$m_{I,1}$	...	$m_{I,j}$	...	$m_{I,J}$	$m_{I,\bullet}$
Totaux	$m_{\bullet,1}$	...	$m_{\bullet,j}$	...	$m_{\bullet,J}$	$m_{\bullet,\bullet} = n$

où $m_{i,j}$ correspond au nombre d'individus observés ayant la modalité i pour X et la modalité j pour Y . La notation $m_{i,\bullet}$ correspond à $\sum_{j=1}^J m_{i,j}$ et la notation $m_{\bullet,j}$ correspond à $\sum_{i=1}^I m_{i,j}$. Le principe du test consiste à comparer les effectifs tels que nous les avons, à la répartition que nous aurions si les variables étaient indépendantes. Dans ce cas, en considérant que les marges ($m_{1,\bullet}, \dots, m_{I,\bullet}, m_{\bullet,1}, \dots, m_{\bullet,J}$) sont fixées, nous pouvons calculer cette répartition théorique dans chacun des échantillons. Nous avons alors $c_{i,j} = m_{i,\bullet} m_{\bullet,j} / m_{\bullet,\bullet}$.

► Procédure de test

Hypothèses testées

$\mathcal{H}_0$: Les deux caractères X et Y sont indépendants

contre

$\mathcal{H}_1$: Les deux caractères X et Y ne sont pas indépendants.

Conditions d'application du test

Considérons un échantillon formé de réalisations indépendantes du couple de variables aléatoires (X, Y) et de taille $n = m_{\bullet,\bullet}$. Les **effectifs théoriques** $c_{i,j}$ et l'**effectif total de l'échantillon** $m_{\bullet,\bullet}$ doivent vérifier les inégalités :

$$c_{i,j} \geq 5 \quad \text{et} \quad m_{\bullet,\bullet} \geq 50.$$

Lorsque les conditions ne sont pas remplies, il existe des corrections, par exemple celle de Yates, ou le test exact de Fisher, voir le *Maxi Fiches de Statistique* de F. et M. Bertrand, Dunod 2011.

S'il y a plus de deux modalités, nous pouvons essayer d'en regrouper si cela est possible, c'est-à-dire si cela a un sens.

Statistique du test

Si l'hypothèse nulle $\mathcal{H}_0$ est vraie et lorsque les conditions d'application du test sont remplies, $\chi^2(obs) = \sum_{i=1}^I \sum_{j=1}^J \frac{(m_{i,j} - c_{i,j})^2}{c_{i,j}}$ est une réalisation d'une variable aléatoire qui suit approximativement la loi du Khi-deux à $(I - 1) \times (J - 1)$ degrés de liberté.

Règle de décision et conclusion du test

La valeur critique du test, notée $c_{1-\alpha}$, est lue dans une table de la loi du Khi-deux. Si la valeur de la statistique calculée sur l'échantillon, notée $\chi^2(obs)$, est supérieure ou égale à $c_{1-\alpha}$, le test est significatif. Nous rejetons l'hypothèse nulle $\mathcal{H}_0$ et nous décidons que l'hypothèse alternative $\mathcal{H}_1$ est vraie avec un risque de première espèce α . Si la valeur de la statistique calculée sur l'échantillon, notée $\chi^2(obs)$, est inférieure à $c_{1-\alpha}$, le test n'est pas significatif. Nous conservons l'hypothèse nulle $\mathcal{H}_0$ avec un risque de deuxième espèce β à déterminer.

Application

Les incontournables boules

Nous sélectionnons 160 boules dans une urne à quatre couleurs dont nous ne connaissons pas la répartition. À partir de cet échantillon, nous voulons savoir s'il est possible que la répartition des couleurs dans l'urne soit de 9/16 de boules noires, 3/16 de boules rouges, 3/16 de boules jaunes et 1/16 de boules vertes. Les données ont été reportées dans le tableau suivant :

Couleurs	Noir	Rouge	Jaune	Vert	Totaux
Proportions théoriques	$\frac{9}{16}$	$\frac{3}{16}$	$\frac{3}{16}$	$\frac{1}{16}$	1
Effectifs observés	100	18	24	18	160

Est-il plausible, au seuil $\alpha = 5\%$, que l'échantillon (100, 18, 24, 18) provienne d'une urne à quatre couleurs réparties suivant les proportions (9/16, 3/16, 3/16, 1/16) ?

Pour répondre à la question posée, nous allons calculer quels seraient les effectifs si la répartition théorique était respectée. Ainsi avec les 160 boules, si la répartition théorique était respectée nous aurions $160 \times 9/16 = 90$ boules noires. Nous obtenons les résultats suivants :

Couleurs	Noir	Rouge	Jaune	Vert	Total
Effectifs théoriques	90	30	30	10	160

Les effectifs théoriques sont tous supérieurs à 5 et l'effectif total est supérieur à 50. Les conditions d'application du test d'adéquation du Khi-deux sont donc bien remplies. Nous mesurons ensuite la distance entre les effectifs observés et les effectifs théoriques à l'aide de la statistique du Khi-deux :

$$\chi^2(obs) = \frac{(100 - 90)^2}{90} + \frac{(18 - 30)^2}{30} + \frac{(24 - 30)^2}{30} + \frac{(18 - 10)^2}{10} = \frac{10}{9} + \frac{24}{5} + \frac{6}{5} + \frac{32}{5} = \frac{608}{45}.$$

Si l'hypothèse nulle $\mathcal{H}_0$ est vraie, alors $\chi^2(obs)$ est une réalisation d'une variable aléatoire qui suit la loi du χ^2 à $(4 - 1)$ degrés de liberté. Nous obtenons une valeur $\chi^2(obs) = 13,5111$. Pour $\alpha = 5\%$, la valeur critique de la loi du χ^2 à 3 degrés de liberté vaut $c_{0,95} = 7,8147$ et par conséquent $\chi^2(obs) > c_{0,95}$. Le test est significatif au seuil $\alpha = 5\%$. La répartition observée nous permet d'affirmer que la répartition des couleurs au sein de l'urne est significativement différente de la répartition théorique introduite ci-dessus avec un risque d'erreur de première espèce égal à 5 %.

Au seuil $\alpha = 5 \%$, les données suivantes, recueillies par Essenberg, Science 1952, suffisent-elles à montrer un lien entre l'apparition d'une tumeur et l'exposition à la fumée de cigarettes ?

Groupe	Tumeur présente	Tumeur absente	Totaux
Contrôle	19	13	32
Traitement	21	2	23

Pour étudier l'existence de ce lien, procédons à un test du χ^2 . Les dénombrements attendus sont $c_{1,1} = 16,73$, $c_{1,2} = 6,27$, $c_{2,1} = 23,27$ et $c_{2,2} = 8,73$. Les conditions d'application du test sont remplies. La valeur critique $c_{0,95}$ est le quantile d'ordre 0,95 d'une loi du Khi-deux à $(2 - 1)(2 - 1) = 1$ degré de liberté. $\chi^2(obs) = 1,091 + 2,910 + 0,784 + 2,092 = 6,878 > 3,8415 = c_{0,95}$. Le test est significatif au seuil $\alpha = 5 \%$: nous décidons, avec un risque d'erreur de première espèce égal à 5 %, qu'il y a un lien significatif entre l'apparition d'une tumeur et l'exposition à la fumée de cigarettes chez les souris.

Fiche 121

Pour s'entraîner (solutions p. 555)

117.1 En reprenant les données de l'exercice 109.1, tester, au risque $\alpha = 5 \%$, l'adéquation entre le nombre d'émissions radioactives pendant une seconde et une loi de Poisson.

117.2 

Existe-t-il, au risque $\alpha = 5 \%$, une relation

entre les habitudes tabagiques d'étudiants, sans liens de parenté entre eux, en Arizona et les habitudes de leurs parents ?

Nombre de parents fumeurs	Enfant fumeur	Enfant non fumeur
Deux	400	1380
Un seul	416	1823
Aucun	188	1168

1. Tests de normalité

► Test de Shapiro-Wilk

Hypothèses testées

$\mathcal{H}_0$: La variable aléatoire X suit une loi normale, *i.e.* $\mathcal{L}(X) = \mathcal{N}$

contre

$\mathcal{H}_1$: La variable aléatoire X ne suit pas une loi normale, *i.e.* $\mathcal{L}(X) \neq \mathcal{N}$.

Conditions d'application du test

L'échantillon $x_1, \dots, x_n$ doit être formé d'au moins 30 réalisations indépendantes de la variable aléatoire X .

Statistique du test

Considérons $(X_1, \dots, X_n)$ un échantillon aléatoire de loi parente la loi de X . Notons $\mathbf{w}$ le vecteur colonne des espérances des statistiques d'ordre, $\mathbf{w} = (w_1, \dots, w_n)$ où w_i est égal à $\mathbb{E}[X_{(i)}]$, la valeur attendue de la i -ème statistique d'ordre $X_{(i)}$, $\mathbf{V}$ la matrice de variance-covariance des statistiques d'ordre $(X_{(1)}, \dots, X_{(n)})$ et $S_{n,c}^2$ l'estimateur corrigé de la variance de l'échantillon. Posons $B = \sum_{i=1}^n a_{i,n} X_{(i)}$ où

$${}^t\mathbf{a}_n = ({}^t\mathbf{w}\mathbf{V}^{-1}\mathbf{V}^{-1}\mathbf{w})^{-1/2}{}^t\mathbf{w}\mathbf{V}^{-1} \quad \text{et} \quad {}^t\mathbf{a}_n\mathbf{a}_n = 1.$$

La statistique du test de Shapiro-Wilk est définie par :

$$W = \frac{B^2}{(n-1)S_{n,c}^2}.$$

Or, par définition $a_{i,n} = -a_{n-i+1,n}$ et de ce fait, en notant $[n/2]$ la partie entière de $n/2$, nous avons :

$$B = \sum_{i=1}^{[n/2]} a_{n-i+1,n} (X_{(n-i+1)} - X_{(i)}).$$



- Les paramètres de la loi normale μ et σ^2 n'ont pas besoin d'être précisés.
- Une des difficultés théoriques pour le calcul de W est la détermination des éléments de $\mathbf{V}$. Ceux-ci ne sont connus de manière exacte que pour des échantillons de taille inférieure à 20.
- Une difficulté calculatoire pour l'évaluation de W réside dans le fait que tous les coefficients $a_{i,n}$ qui interviennent dans le calcul de B dépendent de la taille de l'échantillon.
- Shapiro et Wilk (1965) ont construit des tables qui contiennent les valeurs exactes des coefficients $a_{i,n}$ pour $1 \leq i \leq [n/2]$ et $2 \leq n \leq 50$ et les valeurs critiques de la statistique W pour $3 \leq n \leq 50$. Ces tables sont reproduites dans la Fiche 121.

Fiche 100

Fiche 121

- Royston (1992) a proposé un algorithme pour calculer une valeur approchée de la statistique W du test de Shapiro-Wilk ainsi qu'une approximation de la p -valeur du test, lorsque celle-ci est inférieure à 0, 1, pour $3 \leq n \leq 5000$. C'est cet algorithme qui est utilisé par les logiciels de statistique pour mettre en œuvre le test de Shapiro-Wilk.
- Les techniques statistiques, introduites aux Fiches 110, 112, 114, 115, 116 et 120, reposent sur une hypothèse de normalité. La Fiche 113 d'introduction générale aux tests souligne qu'un rejet de l'hypothèse nulle $\mathcal{H}_0$ au seuil de $\alpha = 5\%$ est associé à un risque d'erreur de première espèce qui vaut au maximum $\alpha = 5\%$ quelle que soit la taille de l'échantillon utilisé pour faire le test. Par contre, conserver l'hypothèse nulle $\mathcal{H}_0$ au seuil de $\alpha = 5\%$ est associé à un risque d'erreur de deuxième espèce qui est difficile à évaluer dans le cas du test de Shapiro-Wilk mais qui diminue lorsque l'effectif de l'échantillon augmente. Plusieurs études par simulations ont montré qu'il faut utiliser un effectif supérieur ou égal à 30 pour parvenir à limiter l'erreur de deuxième espèce à une valeur raisonnable.

Règle de décision et conclusion du test

Lorsque $3 \leq n \leq 50$, les tables de Shapiro et Wilk permettent de déterminer la valeur critique w_α à utiliser pour le test. Si la valeur observée $W(obs)$ de la statistique du test est inférieure ou égale à la valeur critique w_α , alors le test est significatif au seuil α et nous rejetons l'hypothèse nulle $\mathcal{H}_0$. Dans le cas contraire, nous conservons par défaut l'hypothèse nulle $\mathcal{H}_0$.

Lorsque $n > 50$, nous utilisons la p -valeur calculée par un logiciel de statistique. Si la p -valeur est inférieure ou égale au seuil α , alors le test est significatif et nous rejetons l'hypothèse nulle $\mathcal{H}_0$. Dans le cas contraire, nous conservons par défaut l'hypothèse nulle $\mathcal{H}_0$.

► Test de Shapiro-Francia

Les statistiques d'ordre sont d'autant moins fortement corrélées que la taille de l'échantillon est importante. Shapiro et Francia (1972) ont donc proposé comme approximation de la statistique du test de Shapiro-Wilk, introduite à la Section 1, de considérer que les statistiques d'ordre sont indépendantes. Ainsi il n'est plus nécessaire de calculer la matrice de variance-covariance $\mathbf{V}$.

 Fiche 89

Hypothèses testées

$\mathcal{H}_0$: La variable aléatoire X suit une loi normale, i.e. $\mathcal{L}(X) = \mathcal{N}$

contre

$\mathcal{H}_1$: La variable aléatoire X ne suit pas une loi normale, i.e. $\mathcal{L}(X) \neq \mathcal{N}$.

Conditions d'application du test

L'échantillon $(x_1, \dots, x_n)$ doit être formé d'au moins 30 réalisations indépendantes de la variable aléatoire X .

Statistique du test

$$W' = \frac{(\mathbf{a}^* \mathbf{x})^2}{(n-1)S_{n,c}^2} \quad \text{où} \quad \mathbf{a}^* = \frac{1}{\sqrt{\mathbf{w}\mathbf{w}}} \mathbf{w} \quad \text{et} \quad \mathbf{x}_{(\bullet)} = (x_{(1)}, \dots, x_{(n)}).$$

W' est le carré du coefficient de corrélation de Pearson entre $\mathbf{a}^*$ et $\mathbf{x}_{(\bullet)}$, c'est-à-dire le coefficient de détermination R^2 de la régression linéaire de $\mathbf{x}_{(\bullet)}$ sur $\mathbf{a}^*$.

 Fiche 119

Weisberg et Bingham (1975) ont suggéré d'avoir recours à une approximation introduite par Blom (1958) pour approcher, pour tout $1 \leq i \leq n$, les valeurs des w_i :

$$\forall 1 \leq i \leq n, \quad \widetilde{w}_i = \Phi^{-1} \left(\frac{i - 3/8}{n + 1/4} \right),$$

où Φ est la fonction de répartition de la loi normale centrée et réduite, puis d'utiliser $\widetilde{\mathbf{w}} = (\widetilde{w}_1, \dots, \widetilde{w}_n)$ à la place de $\mathbf{w}$ pour calculer W' . C'est généralement cette version simplifiée du test qui est implémentée dans les logiciels de statistique. Si plusieurs observations sont égales, la valeur i dans la formule précédente est à remplacer par la moyenne des rangs qui auraient dû leur être attribués.

Règle de décision et conclusion du test

Les tables de Shapiro et Francia fournies à la Fiche 121 permettent de déterminer la valeur critique w'_α à utiliser pour le test. Si la valeur observée $W'(obs)$ de la statistique du test est inférieure ou égale à la valeur critique w'_α , alors le test est significatif au seuil α et nous rejetons l'hypothèse nulle $\mathcal{H}_0$. Dans le cas contraire, nous conservons par défaut l'hypothèse nulle $\mathcal{H}_0$.

2. Diagramme quantile-quantile

Il est recommandé d'associer systématiquement au test de normalité un diagramme quantile-quantile basé sur la loi normale construit par un logiciel ou en suivant la procédure décrite dans les compléments en ligne. L'hypothèse de normalité est d'autant plus crédible que les points représentés sont alignés, voir la Figure (a) ci-après obtenue pour un échantillon de loi parente la loi normale centrée et réduite. Ce diagramme permet de mettre en évidence certaines caractéristiques de la loi parente de l'échantillon qui auraient pu amener à un rejet de l'hypothèse nulle $\mathcal{H}_0$ de normalité.

- **Distribution leptokurtique**, $\beta_2 > 3$, caractérisée par des queues de distribution plus longues que la loi normale. Ce défaut de normalité est d'autant plus marqué que la partie inférieure du tracé s'incurve vers le bas et la partie supérieure de celui-ci s'incurve vers le haut, voir la Figure (b) ci-après.

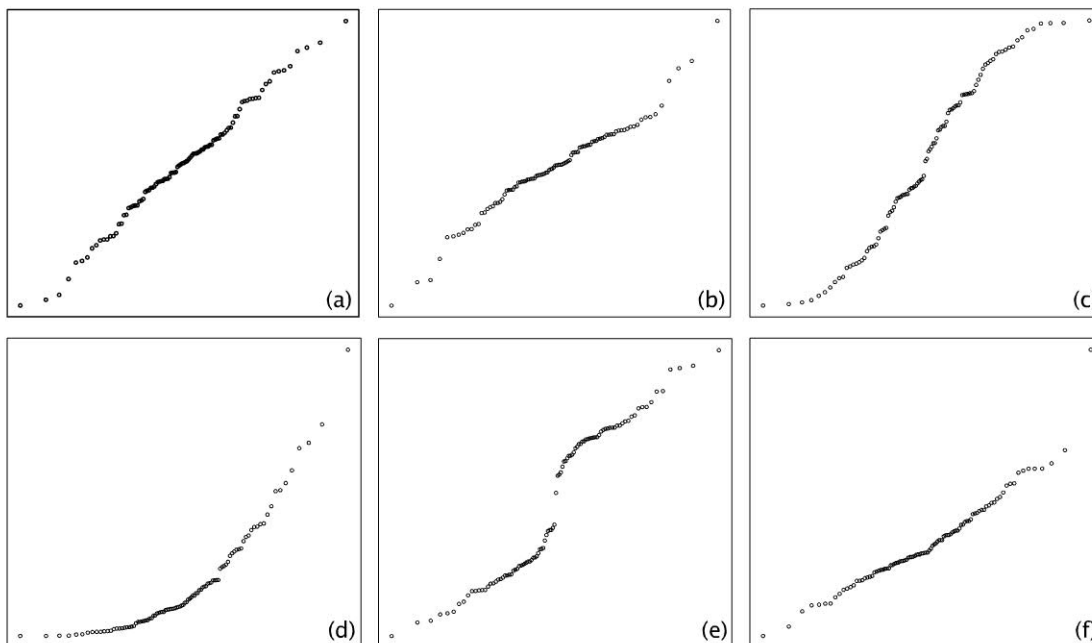
- **Distribution platikurtique**, $\beta_2 < 3$, caractérisée par des queues de distribution plus courtes que la loi normale. Ce défaut de normalité est d'autant plus marqué que le tracé a une courbe sigmoïde, c'est-à-dire en forme de S, voir la Figure (c) ci-après.

- **Distribution asymétrique à droite**, tracé convexe en forme de U, voir la Figure (d) ci-après.

- **Distribution asymétrique à gauche**, tracé concave.

- **Mélange de deux lois normales** de même variance et de moyennes différentes, le tracé se décompose en deux parties approximativement linéaires décalées verticalement l'une par rapport à l'autre, voir la Figure (e) ci-après.

- Échantillon de loi parente normale avec **une valeur extrême**, voir la Figure (f) ci-après.



Application *La vitesse de la lumière*

Reprendre les données de la Fiche 109 concernant la vitesse de la lumière.

1. Mettre en œuvre avec un logiciel (R, SAS ou Maple) le test de Shapiro-Wilk sur l'échantillon.
2. Calculer la réalisation de la statistique W' du test de Shapiro-Francia en ayant recours à l'approximation introduite par Blom. Le test est-il significatif au seuil $\alpha = 5\%$?
3. Au seuil $\alpha = 5\%$, l'échantillon des mesures est-il formé de réalisations d'une loi normale ?

1. Avec un logiciel utilisant la méthode Royston, comme R ou SAS : $W = 0,9881$ et $p\text{-value} = 0,5136$. Avec Maple : $W = 0,9866$ et $p\text{-value} = 0,8545$.

2. Comme $W' = 0,9949 > 0,9878 = w'_{0,05}$, le test n'est pas significatif au seuil $\alpha = 5\%$.

3. Dans tous les cas, nous conservons l'hypothèse nulle $\mathcal{H}_0$ avec un risque de deuxième espèce β . Nous décidons que rien ne s'oppose à l'**hypothèse de normalité** de la variable Vitesse de la Lumière dans l'air. Nous pouvons faire confiance à cette décision car l'effectif de l'échantillon est supérieur à 30 et ainsi nous savons que le risque de deuxième espèce β est maintenu à un niveau acceptable.

Pour s'entraîner (solutions p. 556)

118.1 Appliquer à l'échantillon (0,3780 ; 0,7249 ; 0,1036 ; 0,2905 ; 0,8096 ; 0,8839 ; 0,6986 ; 0,3454 ; 0,9231 ; 0,8785 ; 0,1202 ; 0,8958 ; 0,9528 ; 0,0205 ; 0,1166) les tests de normalité Shapiro-Wilk et Shapiro-Francia. Que concluez-vous au risque $\alpha = 5\%$?

Régression linéaire simple

1. Ajustement linéaire en statistique descriptive

► La méthode des moindres carrés

Nous souhaitons ajuster la distribution du couple de caractères quantitatifs (X, Y) à une fonction affine $Y = b_0 + b_1X$, où b_0 et b_1 sont deux nombres réels à déterminer et nous supposons connaître toutes les valeurs la population, le cas d'un échantillon est traité dans la suite. Cet ajustement s'appelle la **régression linéaire de Y en X** . Soit $(X_1; Y_1), \dots, (X_N; Y_N)$ les données brutes. Le couple (β_0, β_1) va être déterminé de manière à ce que la fonction définie ci-dessous par :

$$d(b_0, b_1) = \sum_{k=1}^N (Y_k - b_0 - b_1 X_k)^2$$

ait un minimum global sur $\mathbb{R}^2$ en (β_0, β_1) . C'est la raison pour laquelle cette méthode s'appelle aussi la **méthode des moindres carrés ordinaires**. Le minimum de cette fonction d positive et différentiable sur $\mathbb{R}^2$ est réalisé en un unique point critique, c'est-à-dire en un point où les deux dérivées partielles de cette fonction sont nulles.



1. $(Y_k - \beta_0 - \beta_1 X_k)^2$ représente le carré de la distance verticale du point (X_k, Y_k) à la droite considérée comme la meilleure.
2. b_1 est le **coefficient directeur** de la droite et b_0 son **ordonnée à l'origine**.
3. La fonction d a pour variables b_0 et b_1 . Les dérivées partielles seront ainsi calculées suivant ces deux variables et les équations définissant le point critique résolues par rapport à ces deux variables. Ici les inconnues ne sont pas les variables X_k et Y_k mais les variables b_0 et b_1 .

► Droite de régression de Y par rapport à X

Définition 119.1

La **droite de régression de Y par rapport à X** est la droite d'équation $y = b_0 + b_1x$ qui rend la fonction $d(b_0, b_1)$ minimale. Les valeurs des nombres réels β_1 et β_0 en sont déduites :

$$\beta_1 = \frac{\mathcal{C}ov(X, Y)}{\sigma^2(X)} = \frac{\mathcal{R}(X, Y)\sigma(Y)}{\sigma(X)} \quad \text{et} \quad \beta_0 = \mu(Y) - \beta_1\mu(X),$$

où $\mathcal{C}ov(X, Y)$ est la covariance de X et de Y , $\sigma^2(X)$ la variance de X , $\mathcal{R}(X, Y)$ le coefficient de corrélation linéaire théorique entre X et Y et $\sigma(Y)$ l'écart-type de Y , $\mu(X)$ et $\mu(Y)$ les moyennes respectives de X et de Y .



1. La droite de régression obtenue par la méthode des moindres carrés, notée D et d'équation $y = \beta_0 + \beta_1x$, passe par le **centre de masse** $(\mu(X), \mu(Y))$ du nuage de points $(X_1, Y_1), \dots, (X_N, Y_N)$.
2. En échangeant les rôles des caractères X et Y , nous obtenons la régression linéaire de X en Y . En général les deux droites de régression sont distinctes.
3. Faites attention à la définition de la covariance qui intervient dans la définition de β_1 . En effet, la série statistique n'est pas donnée par un tableau à double entrée mais par les points relevés et numérotés avec un seul indice k .



Fiches 101
et 105

► Décomposition de la variation et coefficient de détermination

Le but d'un ajustement linéaire est d'expliquer une **partie de la variation** de la variable Y qui est appelée la **variable expliquée** ou **à expliquer** du fait de sa dépendance linéaire à la variable X , qui est appelée la **variable explicative**.

- Toute variable admet une certaine variation, généralement mesurée en terme de variance, c'est-à-dire qu'elle varie d'un individu à l'autre. Pour mesurer la variation de la variable Y , il est d'usage de considérer les différences entre les observations Y_k et leur moyenne $\mu(Y)$. La somme des carrés $\sum_{k=1}^N (Y_k - \mu(Y))^2$ est appelée la **somme des carrés totale** ou **variation totale** et est notée SC_{tot} .
- Si la variable Y dépend de X , et que nous la mesurons sur des individus avec différentes valeurs de X , nous observerons une variation en conséquence. Ici nous ne mesurons que la **variation expliquée** par la régression linéaire de Y en X obtenue par la méthode des moindres carrés. Elle est évaluée par la somme des carrés $\sum_{k=1}^N (\hat{Y}_k - \mu(Y))^2$ qui est appelée la **somme des carrés due à la régression** ou **variation expliquée par le modèle** et est notée SC_{reg} .

$\hat{Y}_k$ représente la **valeur prédite** par cet ajustement pour $X = X_k$, c'est-à-dire $\hat{Y}_k = \beta_0 + \beta_1 X_k$.

- Généralement, lorsque nous mesurons Y sur des individus avec une même valeur de X , nous observons encore une certaine variation. Il s'agit de la **variation inexpliquée** par la régression linéaire de Y en X obtenue par la méthode des moindres carrés.

La somme des carrés $\sum_{k=1}^N (Y_k - \hat{Y}_k)^2$ est appelée la **somme des carrés des résidus** ou la **variation résiduelle** et est notée SC_{res} .

La formule de décomposition de la variance correspond à l'égalité suivante :

$$\text{Variation totale} = \text{Variation expliquée} + \text{Variation inexpliquée}.$$

Définition 119.2

$(\hat{Y}_k - \mu(Y))$ représente la différence expliquée par l'ajustement. $(Y_k - \hat{Y}_k)$ représente la différence inexpliquée par l'ajustement ou le **résidu de l'ajustement**.

Théorème 119.1

La décomposition de la variation se traduit mathématiquement par :

$$\sum_{k=1}^N (Y_k - \mu(Y))^2 = \sum_{k=1}^N (\hat{Y}_k - \mu(Y))^2 + \sum_{k=1}^N (Y_k - \hat{Y}_k)^2.$$

Définition 119.3

Le **coefficient de détermination**, noté R^2 , représente le pourcentage de la variation totale qui est expliquée par l'ajustement et est défini par :

$$R^2 = \frac{\text{Variation expliquée}}{\text{Variation totale}} = \frac{SC_{reg}}{SC_{tot}}.$$

1. Le coefficient de détermination R^2 prend ses valeurs entre 0 et 1.
2. Plus le coefficient de détermination R^2 est proche de 1, plus les données sont alignées sur la droite de régression.



Les caractéristiques de la régression sont résumées dans le **tableau de l'analyse de la variance**.

Sources de Variation	Sommes des carrés	Degrés de liberté	Carrés moyens
Expliquée par la régression	SC_{reg}	1	$CM_{reg} = \frac{SC_{reg}}{1}$
Résiduelle	SC_{res}	$N - 2$	$CM_{res} = \frac{SC_{res}}{N - 2}$
Totale	SC_{tot}	$N - 1$	

2. Estimation ponctuelle des paramètres de la droite de régression

► Modèle gaussien de régression linéaire simple

Il est rare d'avoir accès à toutes les données de la population pour des raisons faciles à comprendre : manque de temps, impossibilité de faire un recensement ...

Pour chaque individu de l'échantillon de taille n , nous observons une valeur x_i du caractère quantitatif continu X et une valeur y_i du caractère quantitatif continu Y . Il peut nous arriver d'observer des valeurs différentes de Y pour plusieurs individus ayant la même valeur observée du caractère X . Derrière cette idée, se trouve le modèle de la régression linéaire simple.

Définition 119.4

La **régression linéaire simple** est liée au modèle suivant :

$$Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \quad \text{où } i = 1, \dots, n.$$

Les deux paramètres β_0 et β_1 sont inconnus mais fixes. Les variables aléatoires ε_i sont les **variables d'erreur** qui sont supposées vérifier des hypothèses, appelées aussi conditions fondamentales, qui sont au nombre de trois :

1. Les variables aléatoires ε_i sont **indépendantes**.
2. Les variables ε_i ont toute la **même variance inconnue σ^2** .
3. Les variables ε_i suivent la **loi normale** centrée et de variance σ^2 .



- Les variables ε_i rendent compte des fluctuations que nous pourrions observer parmi les y_i associées aux mêmes valeurs de x_i .
- Dans la définition ci-dessus, il est mentionné que les paramètres β_0 et β_1 sont inconnus. La variance σ^2 est aussi inconnue. Il faudra aussi l'estimer (cf. Proposition 118.1).

Définition 119.5

La variance σ^2 est appelée la **variance résiduelle** de Y .

► Estimations ponctuelles des paramètres β_1 et β_0

À n données $x_1, \dots, x_n$ de la variable X , l'expérience a associé n données $y_1, \dots, y_n$ de la variable Y . À partir de l'échantillon constitué par les n couples que nous venons d'obtenir, nous pouvons calculer la droite de régression $y = \beta_{0,n} + \beta_{1,n}x$ et le coefficient de corrélation linéaire observé.

Soient $\widehat{\beta}_0$ et $\widehat{\beta}_1$ les variables aléatoires qui prennent les valeurs $\beta_{0,n}$ et $\beta_{1,n}$:

$$\widehat{\beta}_1 = \frac{\sum_{i=1}^n (Y_i - \widehat{\mu}_n(Y))(x_i - \bar{x}_n)}{\sum_{i=1}^n x_i^2 - \bar{x}_n^2} = \frac{\sum_{i=1}^n Y_i x_i - \widehat{\mu}_n(Y) \bar{x}_n}{\sum_{i=1}^n x_i^2 - \bar{x}_n^2} \quad \text{et} \quad \widehat{\beta}_0 = \widehat{\mu}_n(Y) - \widehat{\beta}_1 \bar{x}_n.$$

Les réels $\widehat{\beta}_0(obs) = \beta_{0,n}$ et $\widehat{\beta}_1(obs) = \beta_{1,n}$ sont des estimations de β_0 et β_1 .

Théorème 119.2

$\widehat{\beta}_1$ et $\widehat{\beta}_0$ sont des estimateurs sans biais de β_1 et de β_0 , c'est-à-dire :

$$\mathbb{E}(\widehat{\beta}_1) = \beta_1 \quad \text{et} \quad \mathbb{E}(\widehat{\beta}_0) = \beta_0.$$

► Estimation ponctuelle de la variance résiduelle σ^2

Proposition 119.1

L'estimation de la variance σ^2 peut se faire sans biais par la quantité ci-dessous :

$$s_R^2 = \frac{1}{n-2} \sum_{i=1}^n (y_i - \widehat{\beta}_0 - \widehat{\beta}_1 x_i)^2 = \frac{n}{n-2} (1 - r^2) s^2(y),$$

où $s^2(y) = \frac{1}{n} \sum_{i=1}^n y_i^2 - (\bar{y}_n)^2$, $\bar{y}_n = \frac{1}{n} \sum_{i=1}^n y_i$ et r^2 est la réalisation du coefficient de détermination R^2 évaluée à partir de l'échantillon.

Application

Loi d'Ohm

Lors d'un exercice de travaux pratique, un étudiant doit déterminer si la résistance R d'une résistance qui lui a été confiée est bien conforme au code couleur à quatre anneaux, jaune-vert-noir-marron, qu'elle porte. Ce code couleur signifie que la résistance R de la résistance vaut $450 \, \Omega$ avec une tolérance de 1 %. L'étudiant, qui n'a à sa disposition qu'un ampèremètre et un voltmètre, pense utiliser la loi d'Ohm pour déterminer la valeur de la résistance R . Ayant décidé d'utiliser une régression linéaire simple, il réalise 30 mesures d'intensité et de tension aux bornes de la résistance. Ce nombre de mesures a été choisi afin de pouvoir utiliser le test de Shapiro-Wilk pour vérifier la normalité des erreurs du modèle de regression linéaire simple.

I (A)	U (V)	I (A)	U (V)	I (A)	U (V)	I (A)	U (V)	I (A)	U (V)
0,001	0,479	0,010	4,412	0,020	8,925	0,030	13,609	0,040	18,007
0,002	0,814	0,012	5,345	0,022	9,895	0,032	14,416	0,042	18,830
0,004	1,779	0,014	6,482	0,024	10,645	0,034	15,419	0,044	19,687
0,005	2,212	0,015	6,949	0,025	11,084	0,035	15,774	0,046	20,648
0,006	2,739	0,016	7,099	0,026	11,535	0,036	16,229	0,048	21,453
0,008	3,669	0,018	8,147	0,028	12,562	0,038	17,090	0,050	22,658

1. Les conditions d'utilisation d'un modèle de régression linéaire simple sont-elles remplies ?

2. Donner une estimation ponctuelle des paramètres β_1 et β_0 de la droite de régression, une estimation ponctuelle de la variance résiduelle σ^2 et la valeur du coefficient de détermination R^2 .

1. Le protocole expérimental nous assure de l'indépendance des erreurs ε_i . Le test de Shapiro-Wilk appliqué à l'échantillon formé par les résidus du modèle de régression linéaire simple n'est pas significatif au seuil $\alpha = 5 \%$, $W = 0,9570 > 0,927 = c_{0,05}$: les erreurs sont distribuées normalement. Le Graphique 120.2 ne montre aucun défaut d'égalité des variances puisque les résidus sont sans forme particulière et leur variabilité semble constante. Les trois conditions sont vérifiées.

2. $\widehat{\beta}_0 = -0,0029 \text{ V}$, $\widehat{\beta}_1 = 449,6296 \, \Omega$, $s_R^2 = 0,0103 \text{ V}^2$ et $R^2 = 0,9998$.

 Fiche 120

Intervalles de confiance et tests en régression linéaire simple

Une définition de $\widehat{\beta}_0$, $\widehat{\beta}_1$, $\beta_{0,n}$, $\beta_{1,n}$, $\widehat{\beta}_0(obs)$, $\widehat{\beta}_1(obs)$, et s_R^2 est donnée Fiche 119. Nous notons $\bar{x}_n$ la moyenne de l'échantillon des x_i et $s^2(x)$ sa variance.

 Fiche 119

1. Intervalles de confiance

► Pour la pente β_1

Théorème 120.1

La variance de $\widehat{\beta}_1$ est égale à $\frac{\sigma^2}{ns^2(x)}$. Elle est peut être estimée par $s_{\widehat{\beta}_1}^2 = \frac{s_R^2}{ns^2(x)}$. De plus, sous la condition que les variables aléatoires ε_i sont indépendantes et suivent la loi normale centrée et de variance σ^2 , la variable aléatoire $T_{\widehat{\beta}_1, n-2} = \frac{\widehat{\beta}_1 - \beta_1}{s_{\widehat{\beta}_1}}$ suit la loi de Student $t(n-2)$.



Pour vérifier la condition d'application, il nous faudra réaliser un test de normalité, par exemple le test de Shapiro-Wilk, sur les résidus, ainsi qu'évaluer graphiquement l'homogénéité des variances. Une absence d'homogénéité se traduit par une variabilité non constante des résidus, par exemple des résidus en forme de trompette.

 Fiche 118

De ce théorème, nous déduisons la proposition suivante qui nous donne un intervalle de confiance bilatéral pour la pente β_1 . Le cas unilatéral s'en déduit facilement.

Proposition 120.1

Une réalisation de l'intervalle de confiance bilatéral pour la pente théorique β_1 au niveau de confiance $1 - \alpha$ est égale à

$$\left[\widehat{\beta}_1(obs) - t_{n-2; 1-\alpha/2} \times s_{\widehat{\beta}_1}(obs); \widehat{\beta}_1(obs) + t_{n-2; 1-\alpha/2} \times s_{\widehat{\beta}_1}(obs) \right],$$

où $t_{n-2; 1-\alpha/2}$ est lu dans la table de Student à $n-2$ degrés de liberté.

► Pour l'ordonnée à l'origine β_0

Théorème 120.2

La variance de $\widehat{\beta}_0$ est égale à $\sigma^2 \left(\frac{1}{n} + \frac{\bar{x}_n^2}{ns^2(x)} \right)$. Elle est peut être estimée par $s_{\widehat{\beta}_0}^2 = s_R^2 \left(\frac{1}{n} + \frac{\bar{x}_n^2}{ns^2(x)} \right)$. De plus, sous la condition que les variables aléatoires ε_i sont indépendantes et suivent la loi normale centrée et de variance σ^2 , la variable aléatoire $T_{\widehat{\beta}_0, n-2} = \frac{\widehat{\beta}_0 - \beta_0}{s_{\widehat{\beta}_0}}$ suit la loi de Student $t(n-2)$.



Pour vérifier la condition d'application, il nous faudra réaliser un test de normalité, par exemple le test de Shapiro-Wilk, sur les résidus, ainsi qu'évaluer graphiquement l'homogénéité des variances. Une absence d'homogénéité se traduit par une variabilité non constante des résidus, par exemple des résidus en forme de trompette.

 Fiche 118

De ce théorème, nous déduisons la proposition suivante qui nous donne un intervalle de confiance bilatéral pour l'ordonnée à l'origine β_0 . Le cas unilatéral s'en déduit facilement.

Proposition 120.2

Une réalisation de l'intervalle de confiance bilatéral pour l'ordonnée à l'origine β_0 au niveau de confiance $1 - \alpha$ est égale à

$$\left[\widehat{\beta}_0(obs) - t_{n-2;1-\alpha/2} \times s_{\widehat{\beta}_0}(obs); \widehat{\beta}_0(obs) + t_{n-2;1-\alpha/2} \times s_{\widehat{\beta}_0}(obs) \right],$$

où $t_{n-2;1-\alpha/2}$ est lu dans la table de Student à $n - 2$ degrés de liberté.

► Pour la valeur moyenne de Y pour x fixé

Il est possible de construire un intervalle de confiance bilatéral de la valeur moyenne de Y sachant que $X = x_0$. L'estimation ponctuelle pour cette valeur de x_0 est alors égale à $\widehat{y}_0 = \widehat{\beta}_0(obs) + \widehat{\beta}_1(obs)x_0$. Le cas unilatéral s'en déduit facilement.

Proposition 120.3

Sous la condition que les variables aléatoires ε_i sont indépendantes et suivent la loi normale centrée et de variance σ^2 , au risque α , une réalisation de l'intervalle de confiance bilatéral de la valeur moyenne prise par la variable Y lorsque $X = x_0$ est égale à

$$\left[\widehat{y}_0 - t_{n-2;1-\alpha/2} \sqrt{s_R^2 \left(\frac{1}{n} + \frac{(x_0 - \bar{x}_n)^2}{ns^2(x)} \right)}; \widehat{y}_0 + t_{n-2;1-\alpha/2} \sqrt{s_R^2 \left(\frac{1}{n} + \frac{(x_0 - \bar{x}_n)^2}{ns^2(x)} \right)} \right].$$

► Pour une valeur individuelle de Y pour x fixé

L'ajustement affine peut servir à prévoir une valeur attendue pour la variable Y quand nous fixons $X = x_0$. L'estimation ponctuelle de cette valeur est alors égale à $\widehat{y}_0 = \widehat{\beta}_0(obs) + \widehat{\beta}_1(obs)x_0$.



L'utilisation d'une valeur estimée $\widehat{y}_0$ n'est justifiée que si R^2 est proche de 1.

Proposition 120.4

Sous la condition que les variables aléatoires ε_i sont indépendantes et suivent la loi normale centrée et de variance σ^2 , au risque α , une réalisation de l'intervalle de confiance bilatéral de la valeur prise par la variable Y est égale à

$$\left[\widehat{y}_0 - t_{n-2;1-\alpha/2} \sqrt{s_R^2 \left(1 + \frac{1}{n} + \frac{(x_0 - \bar{x}_n)^2}{ns^2(x)} \right)}; \widehat{y}_0 + t_{n-2;1-\alpha/2} \sqrt{s_R^2 \left(1 + \frac{1}{n} + \frac{(x_0 - \bar{x}_n)^2}{ns^2(x)} \right)} \right].$$

2. Tests sur les paramètres

Les tests unilatéraux se déduisent aisément des tests bilatéraux présentés ci-dessous. Seul le cas où la valeur de référence est nulle est présenté. Mais ce cas se généralise très facilement.

► Comparaison de la pente β_1

Hypothèses testées

$$\mathcal{H}_0 : \beta_1 = 0$$

contre

$$\mathcal{H}_1 : \beta_1 \neq 0.$$

Conditions d'application du test

Les variables aléatoires ε_i sont indépendantes et suivent la loi normale centrée et de variance σ^2 . Pour vérifier la condition d'application, il nous faudra réaliser un test de normalité, par exemple le test de Shapiro-Wilk, sur les résidus, ainsi qu'évaluer graphiquement l'homogénéité des variances. Une absence d'homogénéité se traduit par une variabilité non constante des résidus, par exemple des résidus en forme de trompette.

 Fiche 118

Statistique du test

Si l'hypothèse nulle est vraie et lorsque les conditions d'applications du test sont remplies, la variable aléatoire $T_{\widehat{\beta}_1, n-2} = \frac{\widehat{\beta}_1}{s_{\widehat{\beta}_1}}$ suit la loi de Student $t(n-2)$.

Règle de décision et conclusion du test

La valeur critique du test, notée $c_{1-\alpha/2}$, est lue dans une table de la loi de Student. Si la valeur absolue de la valeur de la statistique calculée sur l'échantillon, notée $T_{\widehat{\beta}_1, n-2}(obs) = \widehat{\beta}_1 / s_{\widehat{\beta}_1}(obs)$, est supérieure ou égale à $c_{1-\alpha/2}$, alors le test est significatif. Nous rejetons $\mathcal{H}_0$ et nous décidons que $\mathcal{H}_1$ est vraie avec un risque de première espèce α . Si la valeur absolue de la valeur de la statistique calculée sur l'échantillon, notée $T_{\widehat{\beta}_1, n-2}(obs) = \widehat{\beta}_1 / s_{\widehat{\beta}_1}(obs)$, est strictement inférieure à $c_{1-\alpha/2}$, alors le test n'est pas significatif. Nous conservons $\mathcal{H}_0$ avec un risque de deuxième espèce β .

► Comparaison de l'ordonnée à l'origine β_0

Hypothèses testées

$$\mathcal{H}_0 : \beta_0 = 0$$

contre

$$\mathcal{H}_1 : \beta_0 \neq 0.$$

Conditions d'application du test

Les variables aléatoires ε_i sont indépendantes et suivent la loi normale centrée et de variance σ^2 . Pour vérifier la condition d'application, il nous faudra réaliser un test de normalité, par exemple le test de Shapiro-Wilk, sur les résidus, ainsi qu'évaluer graphiquement l'homogénéité des variances. Une absence d'homogénéité se traduit par une variabilité non constante des résidus, par exemple des résidus en forme de trompette.

 Fiche 118

Statistique du test

Si l'hypothèse nulle est vraie et lorsque les conditions d'applications du test sont remplies, la variable aléatoire $T_{\hat{\beta}_0, n-2} = \frac{\hat{\beta}_0}{s_{\hat{\beta}_0}}$ suit la loi de Student $t(n-2)$.

Règle de décision et conclusion du test

La valeur critique du test, notée $c_{1-\alpha/2}$, est lue dans une table de la loi de Student. Si la valeur absolue de la valeur de la statistique calculée sur l'échantillon, notée $T_{\hat{\beta}_0, n-2}(obs) = \hat{\beta}_0 / s_{\hat{\beta}_0}(obs)$, est supérieure ou égale à $c_{1-\alpha/2}$, alors le test est significatif. Nous rejetons $\mathcal{H}_0$ et nous décidons que $\mathcal{H}_1$ est vraie avec un risque de première espèce α . Si la valeur absolue de la valeur de la statistique calculée sur l'échantillon, notée $T_{\hat{\beta}_0, n-2}(obs) = \hat{\beta}_0 / s_{\hat{\beta}_0}(obs)$, est strictement inférieure à $c_{1-\alpha/2}$, alors le test n'est pas significatif. Nous conservons $\mathcal{H}_0$ avec un risque de deuxième espèce β .

Application Loi d'Ohm

Reprendre les données de la Fiche 119.

1. Donner une estimation par intervalle de confiance des paramètres β_0 et β_1 avec une confiance de 95 %.

2. Avec une confiance de 95%, les indications présentes sur la résistance concordent-elles avec les résultats expérimentaux obtenus par l'étudiant ?

1. Pour $\alpha = 5\%$, une réalisation de l'intervalle de confiance pour l'ordonnée à l'origine β_0 est $]-0,0029 - 2,0484 \times 0,0360 \text{ V}; -0,0029 + 2,0484 \times 0,0360 \text{ V}[=]-0,0766 \text{ V}; 0,0708 \text{ V}[$. Pour $\alpha = 5\%$, une réalisation de l'intervalle de confiance pour la pente théorique β_1 est $449,6296 - 2,0484 \times 1,2685 \Omega; 449,6296 + 2,0484 \times 1,2685 \Omega [=]447,0312 \Omega; 452,2280 \Omega[$.

2. La résistance vaut 450Ω avec une tolérance de 1%. La réalisation précédente de l'intervalle de confiance est incluse dans ces limites de tolérance : $]447,0312 \Omega; 452,2280 \Omega[\subset]445,5 \Omega; 454,5 \Omega[$. Avec une confiance de 95%, les indications présentes sur la résistance concordent avec les résultats expérimentaux obtenus par l'étudiant.

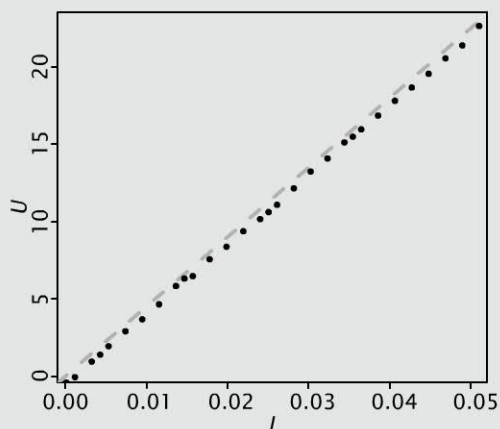


Figure 120.1

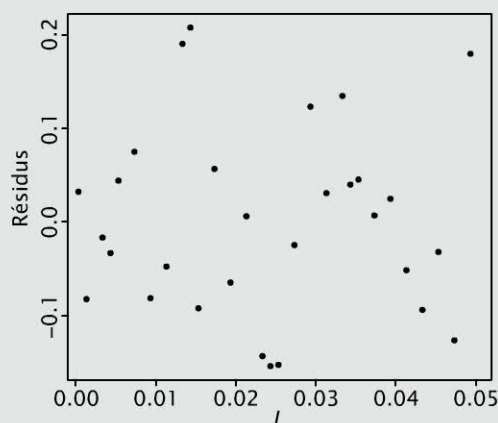


Figure 120.2

Fiche 119

Annexes

1. Fonction de probabilité de la loi de Poisson

$k \setminus \lambda$	1	1,5	2	2,5	3	3,5	4	4,5
0	0,3679	0,2231	0,1353	0,0821	0,0498	0,0302	0,0183	0,0111
1	0,3679	0,3347	0,2707	0,2052	0,1494	0,1057	0,0733	0,0500
2	0,1839	0,2510	0,2707	0,2565	0,2240	0,1850	0,1465	0,1125
3	0,0613	0,1255	0,1804	0,2138	0,2240	0,2158	0,1954	0,1687
4	0,0153	0,0471	0,0902	0,1336	0,1680	0,1888	0,1954	0,1898
5	0,0031	0,0141	0,0361	0,0668	0,1008	0,1322	0,1563	0,1708
6	0,0005	0,0035	0,0120	0,0278	0,0504	0,0771	0,1042	0,1281
7	0,0001	0,0008	0,0034	0,0099	0,0216	0,0385	0,0595	0,0824
8	0,0000	0,0001	0,0009	0,0031	0,0081	0,0169	0,0298	0,0463
9	0,0000	0,0000	0,0002	0,0009	0,0027	0,0066	0,0132	0,0232
10	0,0000	0,0000	0,0000	0,0002	0,0008	0,0023	0,0053	0,0104
11	0,0000	0,0000	0,0000	0,0000	0,0002	0,0007	0,0019	0,0043
12	0,0000	0,0000	0,0000	0,0000	0,0001	0,0002	0,0006	0,0016
13	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0002	0,0006
14	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0002
15	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001
16	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
17	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000

$k \setminus \lambda$	5	5,5	6	6,5	7	7,5	8	8,5
0	0,0067	0,0041	0,0025	0,0015	0,0009	0,0006	0,0003	0,0002
1	0,0337	0,0225	0,0149	0,0098	0,0064	0,0041	0,0027	0,0017
2	0,0842	0,0618	0,0446	0,0318	0,0223	0,0156	0,0107	0,0074
3	0,1404	0,1133	0,0892	0,0688	0,0521	0,0389	0,0286	0,0208
4	0,1755	0,1558	0,1339	0,1118	0,0912	0,0729	0,0573	0,0443
5	0,1755	0,1714	0,1606	0,1454	0,1277	0,1094	0,0916	0,0752
6	0,1462	0,1571	0,1606	0,1575	0,1490	0,1367	0,1221	0,1066
7	0,1044	0,1234	0,1377	0,1462	0,1490	0,1465	0,1396	0,1294
8	0,0653	0,0849	0,1033	0,1188	0,1304	0,1373	0,1396	0,1375
9	0,0363	0,0519	0,0688	0,0858	0,1014	0,1144	0,1241	0,1299
10	0,0181	0,0285	0,0413	0,0558	0,0710	0,0858	0,0993	0,1104
11	0,0082	0,0143	0,0225	0,0330	0,0452	0,0585	0,0722	0,0853
12	0,0034	0,0065	0,0113	0,0179	0,0263	0,0366	0,0481	0,0604
13	0,0013	0,0028	0,0052	0,0089	0,0142	0,0211	0,0296	0,0395
14	0,0005	0,0011	0,0022	0,0041	0,0071	0,0113	0,0169	0,0240
15	0,0002	0,0004	0,0009	0,0018	0,0033	0,0057	0,0090	0,0136
16	0,0000	0,0001	0,0003	0,0007	0,0014	0,0026	0,0045	0,0072
17	0,0000	0,0000	0,0001	0,0003	0,0006	0,0012	0,0021	0,0036

Probabilité $\mathbb{P}(X = k)$ lorsque X est une variable aléatoire qui suit la loi de Poisson de paramètre λ .

2. Fonction de répartition de la loi normale centrée et réduite

	0	,01	,02	,03	,04	,05	,06	,07	,08	,09
0	,5000	,5040	,5080	,5120	,5160	,5199	,5239	,5279	,5319	,5359
,1	,5398	,5438	,5478	,5517	,5557	,5596	,5636	,5675	,5714	,5753
,2	,5793	,5832	,5871	,5910	,5948	,5987	,6026	,6064	,6103	,6141
,3	,6179	,6217	,6255	,6293	,6331	,6368	,6406	,6443	,6480	,6517
,4	,6554	,6591	,6628	,6664	,6700	,6736	,6772	,6808	,6844	,6879
,5	,6915	,6950	,6985	,7019	,7054	,7088	,7123	,7157	,7190	,7224
,6	,7257	,7291	,7324	,7357	,7389	,7422	,7454	,7486	,7517	,7549
,7	,7580	,7611	,7642	,7673	,7704	,7734	,7764	,7794	,7823	,7852
,8	,7881	,7910	,7939	,7967	,7995	,8023	,8051	,8078	,8106	,8133
,9	,8159	,8186	,8212	,8238	,8264	,8289	,8315	,8340	,8365	,8389
1	,8413	,8438	,8461	,8485	,8508	,8531	,8554	,8577	,8599	,8621
1,1	,8643	,8665	,8686	,8708	,8729	,8749	,8770	,8790	,8810	,8830
1,2	,8849	,8869	,8888	,8907	,8925	,8944	,8962	,8980	,8997	,9015
1,3	,9032	,9049	,9066	,9082	,9099	,9115	,9131	,9147	,9162	,9177
1,4	,9192	,9207	,9222	,9236	,9251	,9265	,9279	,9292	,9306	,9319
1,5	,9332	,9345	,9357	,9370	,9382	,9394	,9406	,9418	,9429	,9441
1,6	,9452	,9463	,9474	,9484	,9495	,9505	,9515	,9525	,9535	,9545
1,7	,9554	,9564	,9573	,9582	,9591	,9599	,9608	,9616	,9625	,9633
1,8	,9641	,9649	,9656	,9664	,9671	,9678	,9686	,9693	,9699	,9706
1,9	,9713	,9719	,9726	,9732	,9738	,9744	,9750	,9756	,9761	,9767
2	,9772	,9778	,9783	,9788	,9793	,9798	,9803	,9808	,9812	,9817
2,1	,9821	,9826	,9830	,9834	,9838	,9842	,9846	,9850	,9854	,9857
2,2	,9861	,9864	,9868	,9871	,9875	,9878	,9881	,9884	,9887	,9890
2,3	,9893	,9896	,9898	,9901	,9904	,9906	,9909	,9911	,9913	,9916
2,4	,9918	,9920	,9922	,9925	,9927	,9929	,9931	,9932	,9934	,9936
2,5	,9938	,9940	,9941	,9943	,9945	,9946	,9948	,9949	,9951	,9952
2,6	,9953	,9955	,9956	,9957	,9959	,9960	,9961	,9962	,9963	,9964
2,7	,9965	,9966	,9967	,9968	,9969	,9970	,9971	,9972	,9973	,9974
2,8	,9974	,9975	,9976	,9977	,9977	,9978	,9979	,9979	,9980	,9981
2,9	,9981	,9982	,9982	,9983	,9984	,9984	,9985	,9985	,9986	,9986
3	,9987	,9987	,9987	,9988	,9988	,9989	,9989	,9989	,9990	,9990

Exemple

$\mathbb{P}[X \leq 1,96] = 0,9750$, lorsque X est une variable aléatoire qui suit une $\mathcal{N}(0; 1)$.

3. Quantiles de la loi normale centrée et réduite

	0	0,001	0,002	0,003	0,004	0,005	0,006	0,007	0,008	0,009
0,50	0,000	0,002	0,005	0,007	0,010	0,012	0,015	0,018	0,020	0,023
0,51	0,025	0,028	0,030	0,033	0,035	0,038	0,040	0,043	0,045	0,048
0,52	0,050	0,053	0,055	0,058	0,060	0,063	0,065	0,068	0,070	0,073
0,53	0,075	0,078	0,080	0,083	0,085	0,088	0,090	0,093	0,095	0,098
0,54	0,100	0,103	0,105	0,108	0,110	0,113	0,116	0,118	0,121	0,123
0,55	0,126	0,128	0,131	0,133	0,136	0,138	0,141	0,143	0,146	0,148
0,56	0,151	0,153	0,156	0,159	0,161	0,164	0,166	0,169	0,171	0,174
0,57	0,176	0,179	0,181	0,184	0,187	0,189	0,192	0,194	0,197	0,199
0,58	0,202	0,204	0,207	0,210	0,212	0,215	0,217	0,220	0,222	0,225
0,59	0,228	0,230	0,233	0,235	0,238	0,240	0,243	0,246	0,248	0,251
0,60	0,253	0,256	0,259	0,261	0,264	0,266	0,269	0,272	0,274	0,277
0,61	0,279	0,282	0,284	0,287	0,290	0,292	0,295	0,298	0,300	0,303
0,62	0,305	0,308	0,311	0,313	0,316	0,319	0,321	0,324	0,327	0,329
0,63	0,332	0,335	0,337	0,340	0,343	0,345	0,348	0,350	0,353	0,356
0,64	0,358	0,361	0,364	0,366	0,369	0,372	0,374	0,377	0,380	0,383
0,65	0,385	0,388	0,391	0,393	0,396	0,399	0,402	0,404	0,407	0,410
0,66	0,412	0,415	0,418	0,421	0,423	0,426	0,429	0,432	0,434	0,437
0,67	0,440	0,443	0,445	0,448	0,451	0,454	0,457	0,459	0,462	0,465
0,68	0,468	0,470	0,473	0,476	0,479	0,482	0,484	0,487	0,490	0,493
0,69	0,496	0,499	0,501	0,504	0,507	0,510	0,513	0,516	0,519	0,521
0,70	0,524	0,527	0,530	0,533	0,536	0,539	0,542	0,545	0,548	0,550
0,71	0,553	0,556	0,559	0,562	0,565	0,568	0,571	0,574	0,577	0,580
0,72	0,583	0,586	0,589	0,592	0,595	0,598	0,601	0,604	0,607	0,610
0,73	0,613	0,616	0,619	0,622	0,625	0,628	0,631	0,634	0,637	0,640
0,74	0,643	0,646	0,649	0,653	0,656	0,659	0,662	0,665	0,668	0,671
0,75	0,674	0,678	0,681	0,684	0,687	0,690	0,694	0,697	0,700	0,703
0,76	0,706	0,710	0,713	0,716	0,719	0,723	0,726	0,729	0,732	0,736
0,77	0,739	0,742	0,745	0,749	0,752	0,755	0,759	0,762	0,765	0,769
0,78	0,772	0,776	0,779	0,782	0,786	0,789	0,793	0,796	0,799	0,803
0,79	0,806	0,810	0,813	0,817	0,820	0,824	0,827	0,831	0,835	0,838
0,80	0,842	0,845	0,849	0,852	0,856	0,860	0,863	0,867	0,871	0,874
0,81	0,878	0,882	0,885	0,889	0,893	0,896	0,900	0,904	0,908	0,912
0,82	0,915	0,919	0,923	0,927	0,931	0,935	0,938	0,942	0,946	0,950
0,83	0,954	0,958	0,962	0,966	0,970	0,974	0,978	0,982	0,986	0,990
0,84	0,995	0,999	1,003	1,007	1,011	1,015	1,019	1,024	1,028	1,032
0,85	1,036	1,041	1,045	1,049	1,054	1,058	1,062	1,067	1,071	1,076
0,86	1,080	1,085	1,089	1,094	1,099	1,103	1,108	1,112	1,117	1,122
0,87	1,126	1,131	1,136	1,141	1,145	1,150	1,155	1,160	1,165	1,170
0,88	1,175	1,180	1,185	1,190	1,195	1,200	1,206	1,211	1,216	1,221
0,89	1,226	1,232	1,237	1,243	1,248	1,254	1,259	1,265	1,270	1,276
0,90	1,282	1,287	1,293	1,299	1,305	1,311	1,316	1,323	1,329	1,335
0,91	1,341	1,347	1,353	1,359	1,366	1,372	1,379	1,385	1,392	1,398
0,92	1,405	1,412	1,419	1,425	1,433	1,440	1,447	1,454	1,461	1,468
0,93	1,476	1,483	1,491	1,498	1,506	1,514	1,522	1,530	1,538	1,546
0,94	1,555	1,563	1,572	1,581	1,589	1,598	1,607	1,616	1,626	1,635
0,95	1,645	1,655	1,665	1,675	1,685	1,695	1,706	1,717	1,728	1,739
0,96	1,751	1,762	1,774	1,787	1,799	1,812	1,825	1,838	1,852	1,866
0,97	1,881	1,896	1,911	1,927	1,943	1,960	1,977	1,995	2,014	2,034
0,98	2,054	2,075	2,097	2,120	2,144	2,170	2,197	2,226	2,257	2,290
0,99	2,326	2,366	2,409	2,457	2,512	2,576	2,652	2,748	2,878	3,090

Exemple

$\mathbb{P}[X \leq 1,96] = 0,9750$, lorsque X est une variable aléatoire qui suit une $\mathcal{N}(0; 1)$.

4. Loi de Student

	0,900	0,950	0,975	0,990	0,995
1	3,0777	6,3138	12,7062	31,8205	63,6567
2	1,8856	2,9200	4,3027	6,9646	9,9248
3	1,6377	2,3534	3,1824	4,5407	5,8409
4	1,5332	2,1318	2,7764	3,7469	4,6041
5	1,4759	2,0150	2,5706	3,3649	4,0321
6	1,4398	1,9432	2,4469	3,1427	3,7074
7	1,4149	1,8946	2,3646	2,9980	3,4995
8	1,3968	1,8595	2,3060	2,8965	3,3554
9	1,3830	1,8331	2,2622	2,8214	3,2498
10	1,3722	1,8125	2,2281	2,7638	3,1693
11	1,3634	1,7959	2,2010	2,7181	3,1058
12	1,3562	1,7823	2,1788	2,6810	3,0545
13	1,3502	1,7709	2,1604	2,6503	3,0123
14	1,3450	1,7613	2,1448	2,6245	2,9768
15	1,3406	1,7531	2,1314	2,6025	2,9467
16	1,3368	1,7459	2,1199	2,5835	2,9208
17	1,3334	1,7396	2,1098	2,5669	2,8982
18	1,3304	1,7341	2,1009	2,5524	2,8784
19	1,3277	1,7291	2,0930	2,5395	2,8609
20	1,3253	1,7247	2,0860	2,5280	2,8453
21	1,3232	1,7207	2,0796	2,5176	2,8314
22	1,3212	1,7171	2,0739	2,5083	2,8188
23	1,3195	1,7139	2,0687	2,4999	2,8073
24	1,3178	1,7109	2,0639	2,4922	2,7969
25	1,3163	1,7081	2,0595	2,4851	2,7874
26	1,3150	1,7056	2,0555	2,4786	2,7787
27	1,3137	1,7033	2,0518	2,4727	2,7707
28	1,3125	1,7011	2,0484	2,4671	2,7633
29	1,3114	1,6991	2,0452	2,4620	2,7564
30	1,3104	1,6973	2,0423	2,4573	2,7500
31	1,3095	1,6955	2,0395	2,4528	2,7440
32	1,3086	1,6939	2,0369	2,4487	2,7385
33	1,3077	1,6924	2,0345	2,4448	2,7333
34	1,3070	1,6909	2,0322	2,4411	2,7284
35	1,3062	1,6896	2,0301	2,4377	2,7238
36	1,3055	1,6883	2,0281	2,4345	2,7195
37	1,3049	1,6871	2,0262	2,4314	2,7154
38	1,3042	1,6860	2,0244	2,4286	2,7116
39	1,3036	1,6849	2,0227	2,4258	2,7079
40	1,3031	1,6839	2,0211	2,4233	2,7045
41	1,3025	1,6829	2,0195	2,4208	2,7012
42	1,3020	1,6820	2,0181	2,4185	2,6981
43	1,3016	1,6811	2,0167	2,4163	2,6951
44	1,3011	1,6802	2,0154	2,4141	2,6923
45	1,3006	1,6794	2,0141	2,4121	2,6896
46	1,3002	1,6787	2,0129	2,4102	2,6870
47	1,2998	1,6779	2,0117	2,4083	2,6846
48	1,2994	1,6772	2,0106	2,4066	2,6822
49	1,2991	1,6766	2,0096	2,4049	2,6800
50	1,2987	1,6759	2,0086	2,4033	2,6778

Exemple

$\mathbb{P}[X \leq 3,1824] = 0,975$, lorsque X est une variable aléatoire qui suit une $t(3)$.

5. Loi du χ^2 , valeurs critiques inférieures

	0,005	0,01	0,025	0,05	0,1
1	0,0000	0,0002	0,0010	0,0039	0,0158
2	0,0100	0,0201	0,0506	0,1026	0,2107
3	0,0717	0,1148	0,2158	0,3518	0,5844
4	0,2070	0,2971	0,4844	0,7107	1,0636
5	0,4117	0,5543	0,8312	1,1455	1,6103
6	0,6757	0,8721	1,2373	1,6354	2,2041
7	0,9893	1,2390	1,6899	2,1673	2,8331
8	1,3444	1,6465	2,1797	2,7326	3,4895
9	1,7349	2,0879	2,7004	3,3251	4,1682
10	2,1559	2,5582	3,2470	3,9403	4,8652
11	2,6032	3,0535	3,8157	4,5748	5,5778
12	3,0738	3,5706	4,4038	5,2260	6,3038
13	3,5650	4,1069	5,0088	5,8919	7,0415
14	4,0747	4,6604	5,6287	6,5706	7,7895
15	4,6009	5,2293	6,2621	7,2609	8,5468
16	5,1422	5,8122	6,9077	7,9616	9,3122
17	5,6972	6,4078	7,5642	8,6718	10,0852
18	6,2648	7,0149	8,2307	9,3905	10,8649
19	6,8440	7,6327	8,9065	10,1170	11,6509
20	7,4338	8,2604	9,5908	10,8508	12,4426
21	8,0337	8,8972	10,2829	11,5913	13,2396
22	8,6427	9,5425	10,9823	12,3380	14,0415
23	9,2604	10,1957	11,6886	13,0905	14,8480
24	9,8862	10,8564	12,4012	13,8484	15,6587
25	10,5197	11,5240	13,1197	14,6114	16,4734
26	11,1602	12,1981	13,8439	15,3792	17,2919
27	11,8076	12,8785	14,5734	16,1514	18,1139
28	12,4613	13,5647	15,3079	16,9279	18,9392
29	13,1211	14,2565	16,0471	17,7084	19,7677
30	13,7867	14,9535	16,7908	18,4927	20,5992

6. Loi du χ^2 , valeurs critiques supérieures

	0,9	0,95	0,975	0,99	0,995
1	2,7055	3,8415	5,0239	6,6349	7,8794
2	4,6052	5,9915	7,3778	9,2103	10,5966
3	6,2514	7,8147	9,3484	11,3449	12,8382
4	7,7794	9,4877	11,1433	13,2767	14,8603
5	9,2364	11,0705	12,8325	15,0863	16,7496
6	10,6446	12,5916	14,4494	16,8119	18,5476
7	12,0170	14,0671	16,0128	18,4753	20,2777
8	13,3616	15,5073	17,5345	20,0902	21,9550
9	14,6837	16,9190	19,0228	21,6660	23,5894
10	15,9872	18,3070	20,4832	23,2093	25,1882
11	17,2750	19,6751	21,9200	24,7250	26,7568
12	18,5493	21,0261	23,3367	26,2170	28,2995
13	19,8119	22,3620	24,7356	27,6882	29,8195
14	21,0641	23,6848	26,1189	29,1412	31,3193
15	22,3071	24,9958	27,4884	30,5779	32,8013
16	23,5418	26,2962	28,8454	31,9999	34,2672
17	24,7690	27,5871	30,1910	33,4087	35,7185
18	25,9894	28,8693	31,5264	34,8053	37,1565
19	27,2036	30,1435	32,8523	36,1909	38,5823
20	28,4120	31,4104	34,1696	37,5662	39,9968
21	29,6151	32,6706	35,4789	38,9322	41,4011
22	30,8133	33,9244	36,7807	40,2894	42,7957
23	32,0069	35,1725	38,0756	41,6384	44,1813
24	33,1962	36,4150	39,3641	42,9798	45,5585
25	34,3816	37,6525	40,6465	44,3141	46,9279
26	35,5632	38,8851	41,9232	45,6417	48,2899
27	36,7412	40,1133	43,1945	46,9629	49,6449
28	37,9159	41,3371	44,4608	48,2782	50,9934
29	39,0875	42,5570	45,7223	49,5879	52,3356
30	40,2560	43,7730	46,9792	50,8922	53,6720

Exemples

1. $\mathbb{P}[X \leq 0,0506] = 0,025$, lorsque X est une variable aléatoire qui suit une $\chi^2(2)$.
2. $\mathbb{P}[X \leq 7,3778] = 0,975$, lorsque X est une variable aléatoire qui suit une $\chi^2(2)$.
3. $\mathbb{P}[X \leq 18,5128] = 0,95 = 1 - \alpha$, lorsque X est une variable aléatoire qui suit une $F(1; 2)$ et avec $\alpha = 5\%$.
4. $\mathbb{P}[X \leq 38,5063] = 0,975 = 1 - \alpha$, lorsque X est une variable aléatoire qui suit une $F(1; 2)$ et avec $\alpha = 2,5\%$.

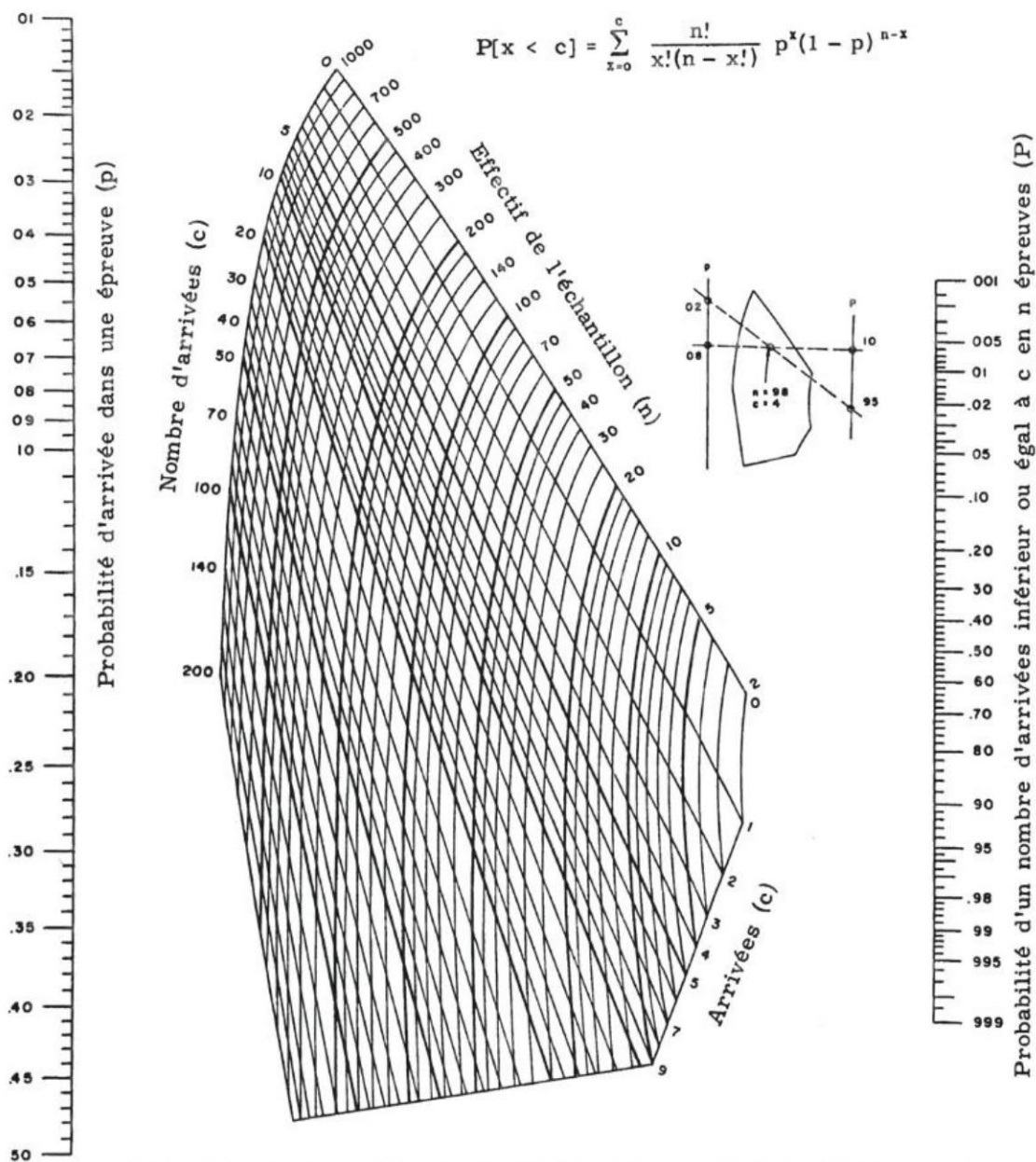
7. Table de la loi de Fisher-Snedecor, $\alpha = 5 \%$

num	den 1	2	3	4	5	6	7	8	9	10
1	161,4476	18,5128	10,1280	7,7086	6,6079	5,9874	5,5914	5,3177	5,1174	4,9646
2	199,5000	19,0000	9,5521	6,9443	5,7861	5,1433	4,7374	4,4590	4,2565	4,1028
3	215,7073	19,1643	9,2766	6,5914	5,4095	4,7571	4,3468	4,0662	3,8625	3,7083
4	224,5832	19,2468	9,1172	6,3882	5,1922	4,5337	4,1203	3,8379	3,6331	3,4780
5	230,1619	19,2964	9,0135	6,2561	5,0503	4,3874	3,9715	3,6875	3,4817	3,3258
6	233,9860	19,3295	8,9406	6,1631	4,9503	4,2839	3,8660	3,5806	3,3738	3,2172
7	236,7684	19,3532	8,8867	6,0942	4,8759	4,2067	3,7870	3,5005	3,2927	3,1355
8	238,8827	19,3710	8,8452	6,0410	4,8183	4,1468	3,7257	3,4381	3,2296	3,0717
9	240,5433	19,3848	8,8123	5,9988	4,7725	4,0990	3,6767	3,3881	3,1789	3,0204
10	241,8817	19,3959	8,7855	5,9644	4,7351	4,0600	3,6365	3,3472	3,1373	2,9782
11	242,9835	19,4050	8,7633	5,9358	4,7040	4,0274	3,6030	3,3130	3,1025	2,9430
12	243,9060	19,4125	8,7446	5,9117	4,6777	3,9999	3,5747	3,2839	3,0729	2,9130
13	244,6898	19,4189	8,7287	5,8911	4,6552	3,9764	3,5503	3,2590	3,0475	2,8872
14	245,3640	19,4244	8,7149	5,8733	4,6358	3,9559	3,5292	3,2374	3,0255	2,8647
15	245,9499	19,4291	8,7029	5,8578	4,6188	3,9381	3,5107	3,2184	3,0061	2,8450
16	246,4639	19,4333	8,6923	5,8441	4,6038	3,9223	3,4944	3,2016	2,9890	2,8276
17	246,9184	19,4370	8,6829	5,8320	4,5904	3,9083	3,4799	3,1867	2,9737	2,8120
18	247,3232	19,4402	8,6745	5,8211	4,5785	3,8957	3,4669	3,1733	2,9600	2,7980
19	247,6861	19,4431	8,6670	5,8114	4,5678	3,8844	3,4551	3,1613	2,9477	2,7854
20	248,0131	19,4458	8,6602	5,8025	4,5581	3,8742	3,4445	3,1503	2,9365	2,7740
21	248,3094	19,4481	8,6540	5,7945	4,5493	3,8649	3,4349	3,1404	2,9263	2,7636
22	248,5791	19,4503	8,6484	5,7872	4,5413	3,8564	3,4260	3,1313	2,9169	2,7541
23	248,8256	19,4523	8,6432	5,7805	4,5339	3,8486	3,4179	3,1229	2,9084	2,7453
24	249,0518	19,4541	8,6385	5,7744	4,5272	3,8415	3,4105	3,1152	2,9005	2,7372
25	249,2601	19,4558	8,6341	5,7687	4,5209	3,8348	3,4036	3,1081	2,8932	2,7298
26	249,4525	19,4573	8,6301	5,7635	4,5151	3,8287	3,3972	3,1015	2,8864	2,7229
27	249,6309	19,4587	8,6263	5,7586	4,5097	3,8230	3,3913	3,0954	2,8801	2,7164
28	249,7966	19,4600	8,6229	5,7541	4,5047	3,8177	3,3858	3,0897	2,8743	2,7104
29	249,9510	19,4613	8,6196	5,7498	4,5001	3,8128	3,3806	3,0844	2,8688	2,7048
30	250,0951	19,4624	8,6166	5,7459	4,4957	3,8082	3,3758	3,0794	2,8637	2,6996
num	den 11	12	13	14	15	16	17	18	19	20
1	4,8443	4,7472	4,6672	4,6001	4,5431	4,4940	4,4513	4,4139	4,3807	4,3512
2	3,9823	3,8853	3,8056	3,7389	3,6823	3,6337	3,5915	3,5546	3,5219	3,4928
3	3,5874	3,4903	3,4105	3,3439	3,2874	3,2389	3,1968	3,1599	3,1274	3,0984
4	3,3567	3,2592	3,1791	3,1122	3,0556	3,0069	2,9647	2,9277	2,8951	2,8661
5	3,2039	3,1059	3,0254	2,9582	2,9013	2,8524	2,8100	2,7729	2,7401	2,7109
6	3,0946	2,9961	2,9153	2,8477	2,7905	2,7413	2,6987	2,6613	2,6283	2,5990
7	3,0123	2,9134	2,8321	2,7642	2,7066	2,6572	2,6143	2,5767	2,5435	2,5140
8	2,9480	2,8486	2,7669	2,6987	2,6408	2,5911	2,5480	2,5102	2,4768	2,4471
9	2,8962	2,7964	2,7144	2,6458	2,5876	2,5377	2,4943	2,4563	2,4227	2,3928
10	2,8536	2,7534	2,6710	2,6022	2,5437	2,4935	2,4499	2,4117	2,3779	2,3479
11	2,8179	2,7173	2,6347	2,5655	2,5068	2,4564	2,4126	2,3742	2,3402	2,3100
12	2,7876	2,6866	2,6037	2,5342	2,4753	2,4247	2,3807	2,3421	2,3080	2,2776
13	2,7614	2,6602	2,5769	2,5073	2,4481	2,3973	2,3531	2,3143	2,2800	2,2495
14	2,7386	2,6371	2,5536	2,4837	2,4244	2,3733	2,3290	2,2900	2,2556	2,2250
15	2,7186	2,6169	2,5331	2,4630	2,4034	2,3522	2,3077	2,2686	2,2341	2,2033
16	2,7009	2,5989	2,5149	2,4446	2,3849	2,3335	2,2888	2,2496	2,2149	2,1840
17	2,6851	2,5828	2,4987	2,4282	2,3683	2,3167	2,2719	2,2325	2,1977	2,1667
18	2,6709	2,5684	2,4841	2,4134	2,3533	2,3016	2,2567	2,2172	2,1823	2,1511
19	2,6581	2,5554	2,4709	2,4000	2,3398	2,2880	2,2429	2,2033	2,1683	2,1370
20	2,6464	2,5436	2,4589	2,3879	2,3275	2,2756	2,2304	2,1906	2,1555	2,1242
21	2,6358	2,5328	2,4479	2,3768	2,3163	2,2642	2,2189	2,1791	2,1438	2,1124
22	2,6261	2,5229	2,4379	2,3667	2,3060	2,2538	2,2084	2,1685	2,1331	2,1016
23	2,6172	2,5139	2,4287	2,3573	2,2966	2,2443	2,1987	2,1587	2,1233	2,0917
24	2,6090	2,5055	2,4202	2,3487	2,2878	2,2354	2,1898	2,1497	2,1141	2,0825
25	2,6014	2,4977	2,4123	2,3407	2,2797	2,2272	2,1815	2,1413	2,1057	2,0739
26	2,5943	2,4905	2,4050	2,3333	2,2722	2,2196	2,1738	2,1335	2,0978	2,0660
27	2,5877	2,4838	2,3982	2,3264	2,2652	2,2125	2,1666	2,1262	2,0905	2,0586
28	2,5816	2,4776	2,3918	2,3199	2,2587	2,2059	2,1599	2,1195	2,0836	2,0517
29	2,5759	2,4718	2,3859	2,3139	2,2525	2,1997	2,1536	2,1131	2,0772	2,0452
30	2,5705	2,4663	2,3803	2,3082	2,2468	2,1938	2,1477	2,1071	2,0712	2,0391

8. Table de la loi de Fisher-Snedecor, $\alpha = 2,5 \%$

num	den 1	2	3	4	5	6	7	8	9	10
1	647,7890	38,5063	17,4434	12,2179	10,0070	8,8131	8,0727	7,5709	7,2093	6,9367
2	799,5000	39,0000	16,0441	10,6491	8,4336	7,2599	6,5415	6,0595	5,7147	5,4564
3	864,1630	39,1655	15,4392	9,9792	7,7636	6,5988	5,8898	5,4160	5,0781	4,8256
4	899,5833	39,2484	15,1010	9,6045	7,3879	6,2272	5,5226	5,0526	4,7181	4,4683
5	921,8479	39,2982	14,8848	9,3645	7,1464	5,9876	5,2852	4,8173	4,4844	4,2361
6	937,1111	39,3315	14,7347	9,1973	6,9777	5,8198	5,1186	4,6517	4,3197	4,0721
7	948,2169	39,3552	14,6244	9,0741	6,8531	5,6955	4,9949	4,5286	4,1970	3,9498
8	956,6562	39,3730	14,5399	8,9796	6,7572	5,5996	4,8993	4,4333	4,1020	3,8549
9	963,2846	39,3869	14,4731	8,9047	6,6811	5,5234	4,8232	4,3572	4,0260	3,7790
10	968,6274	39,3980	14,4189	8,8439	6,6192	5,4613	4,7611	4,2951	3,9639	3,7168
11	973,0252	39,4071	14,3742	8,7935	6,5678	5,4098	4,7095	4,2434	3,9121	3,6649
12	976,7079	39,4146	14,3366	8,7512	6,5245	5,3662	4,6658	4,1997	3,8682	3,6209
13	979,8368	39,4210	14,3045	8,7150	6,4876	5,3290	4,6285	4,1622	3,8306	3,5832
14	982,5278	39,4265	14,2768	8,6838	6,4556	5,2968	4,5961	4,1297	3,7980	3,5504
15	984,8668	39,4313	14,2527	8,6565	6,4277	5,2687	4,5678	4,1012	3,7694	3,5217
16	986,9187	39,4354	14,2315	8,6326	6,4032	5,2439	4,5428	4,0761	3,7441	3,4963
17	988,7331	39,4391	14,2127	8,6113	6,3814	5,2218	4,5206	4,0538	3,7216	3,4737
18	990,3490	39,4424	14,1960	8,5924	6,3619	5,2021	4,5008	4,0338	3,7015	3,4534
19	991,7973	39,4453	14,1810	8,5753	6,3444	5,1844	4,4829	4,0158	3,6833	3,4351
20	993,1028	39,4479	14,1674	8,5599	6,3286	5,1684	4,4667	3,9995	3,6669	3,4185
21	994,2856	39,4503	14,1551	8,5460	6,3142	5,1538	4,4520	3,9846	3,6520	3,4035
22	995,3622	39,4525	14,1438	8,5332	6,3011	5,1406	4,4386	3,9711	3,6383	3,3897
23	996,3462	39,4544	14,1336	8,5216	6,2891	5,1284	4,4263	3,9587	3,6257	3,3770
24	997,2492	39,4562	14,1241	8,5109	6,2780	5,1172	4,4150	3,9472	3,6142	3,3654
25	998,0808	39,4579	14,1155	8,5010	6,2679	5,1069	4,4045	3,9367	3,6035	3,3546
26	998,8490	39,4594	14,1074	8,4919	6,2584	5,0973	4,3949	3,9269	3,5936	3,3446
27	999,5609	39,4609	14,1000	8,4834	6,2497	5,0884	4,3859	3,9178	3,5845	3,3353
28	1000,2225	39,4622	14,0930	8,4755	6,2416	5,0802	4,3775	3,9093	3,5759	3,3267
29	1000,8388	39,4634	14,0866	8,4681	6,2340	5,0724	4,3697	3,9014	3,5679	3,3186
30	1001,4144	39,4646	14,0805	8,4613	6,2269	5,0652	4,3624	3,8940	3,5604	3,3110
num	den 11	12	13	14	15	16	17	18	19	20
1	6,7241	6,5538	6,4143	6,2979	6,1995	6,1151	6,0420	5,9781	5,9216	5,8715
2	5,2559	5,0959	4,9653	4,8567	4,7650	4,6867	4,6189	4,5597	4,5075	4,4613
3	4,6300	4,4742	4,3472	4,2417	4,1528	4,0768	4,0112	3,9539	3,9034	3,8587
4	4,2751	4,1212	3,9959	3,8919	3,8043	3,7294	3,6648	3,6083	3,5587	3,5147
5	4,0440	3,8911	3,7667	3,6634	3,5764	3,5021	3,4379	3,3820	3,3327	3,2891
6	3,8807	3,7283	3,6043	3,5014	3,4147	3,3406	3,2767	3,2209	3,1718	3,1283
7	3,7586	3,6065	3,4827	3,3799	3,2934	3,2194	3,1556	3,0999	3,0509	3,0074
8	3,6638	3,5118	3,3880	3,2853	3,1987	3,1248	3,0610	3,0053	2,9563	2,9128
9	3,5879	3,4358	3,3120	3,2093	3,1227	3,0488	2,9849	2,9291	2,8801	2,8365
10	3,5257	3,3736	3,2497	3,1469	3,0602	2,9862	2,9222	2,8664	2,8172	2,7737
11	3,4737	3,3215	3,1975	3,0946	3,0078	2,9337	2,8696	2,8137	2,7645	2,7209
12	3,4296	3,2773	3,1532	3,0502	2,9633	2,8890	2,8249	2,7689	2,7196	2,6758
13	3,3917	3,2393	3,1150	3,0119	2,9249	2,8506	2,7863	2,7302	2,6808	2,6369
14	3,3588	3,2062	3,0819	2,9786	2,8915	2,8170	2,7526	2,6964	2,6469	2,6030
15	3,3299	3,1772	3,0527	2,9493	2,8621	2,7875	2,7230	2,6667	2,6171	2,5731
16	3,3044	3,1515	3,0269	2,9234	2,8360	2,7614	2,6968	2,6404	2,5907	2,5465
17	3,2816	3,1286	3,0039	2,9003	2,8128	2,7380	2,6733	2,6168	2,5670	2,5228
18	3,2612	3,1081	2,9832	2,8795	2,7919	2,7170	2,6522	2,5956	2,5457	2,5014
19	3,2428	3,0896	2,9646	2,8607	2,7730	2,6980	2,6331	2,5764	2,5265	2,4821
20	3,2261	3,0728	2,9477	2,8437	2,7559	2,6808	2,6158	2,5590	2,5089	2,4645
21	3,2109	3,0575	2,9322	2,8282	2,7403	2,6651	2,6000	2,5431	2,4930	2,4484
22	3,1970	3,0434	2,9181	2,8139	2,7260	2,6507	2,5855	2,5285	2,4783	2,4337
23	3,1843	3,0306	2,9052	2,8009	2,7128	2,6374	2,5721	2,5151	2,4648	2,4201
24	3,1725	3,0187	2,8932	2,7888	2,7006	2,6252	2,5598	2,5027	2,4523	2,4076
25	3,1616	3,0077	2,8821	2,7777	2,6894	2,6138	2,5484	2,4912	2,4408	2,3959
26	3,1516	2,9976	2,8719	2,7673	2,6790	2,6033	2,5378	2,4806	2,4300	2,3851
27	3,1422	2,9881	2,8623	2,7577	2,6692	2,5935	2,5280	2,4706	2,4200	2,3751
28	3,1334	2,9793	2,8534	2,7487	2,6602	2,5844	2,5187	2,4613	2,4107	2,3657
29	3,1253	2,9710	2,8451	2,7403	2,6517	2,5758	2,5101	2,4527	2,4019	2,3569
30	3,1176	2,9633	2,8372	2,7324	2,6437	2,5678	2,5020	2,4445	2,3937	2,3486

9. Nomogramme de Larson



NOMOGRAMME DE LA DISTRIBUTION BINOMIALE CUMULATIVE

Nomogramme de la distribution binomiale cumulative. Revue de Statistique Appliquée, 1967 – vol. XV – N° 2.

10. Statistique W du test de normalité de Shapiro-Wilk : coefficients $a_{i,n}$

Taille de l'échantillon

i	2	3	4	5	6	7	8	
1	0,7071	0,7071	0,6872	0,6646	0,6431	0,6233	0,6052	
2	-	0,0000	0,1677	0,2413	0,2806	0,3031	0,3164	
3	-	-	-	0,0000	0,0875	0,1401	0,1743	
4	-	-	-	-	-	0,0000	0,0561	
i	9	10	11	12	13	14	15	16
1	0,5888	0,5739	0,5601	0,5475	0,5359	0,5251	0,5150	0,5056
2	0,3244	0,3291	0,3315	0,3325	0,3325	0,3318	0,3306	0,3290
3	0,1976	0,2141	0,2260	0,2347	0,2412	0,2460	0,2495	0,2521
4	0,0947	0,1224	0,1429	0,1586	0,1707	0,1802	0,1878	0,1939
5	0,0000	0,0399	0,0695	0,0922	0,1099	0,1240	0,1353	0,1447
6	-	-	0,0000	0,0303	0,0539	0,0727	0,0880	0,1005
7	-	-	-	-	0,0000	0,0240	0,0433	0,0593
8	-	-	-	-	-	-	0,0000	0,0196
i	17	18	19	20	21	22	23	24
1	0,4968	0,4886	0,4808	0,4734	0,4643	0,4590	0,4542	0,4493
2	0,3273	0,3253	0,3232	0,3211	0,3185	0,3156	0,3126	0,3098
3	0,2540	0,2553	0,2561	0,2565	0,2578	0,2571	0,2563	0,2554
4	0,1988	0,2027	0,2059	0,2085	0,2119	0,2131	0,2139	0,2145
5	0,1524	0,1587	0,1641	0,1686	0,1736	0,1764	0,1787	0,1807
6	0,1109	0,1197	0,1271	0,1334	0,1399	0,1443	0,1480	0,1512
7	0,0725	0,0837	0,0932	0,1013	0,1092	0,1150	0,1201	0,1245
8	0,0359	0,0496	0,0612	0,0711	0,0804	0,0878	0,0941	0,0997
9	0,0000	0,0163	0,0303	0,0422	0,0530	0,0618	0,0696	0,0764
10	-	-	0,0000	0,0140	0,0263	0,0368	0,0459	0,0539
11	-	-	-	-	0,0000	0,0122	0,0228	0,0321
12	-	-	-	-	-	-	0,0000	0,0107
i	25	26	27	28	29	30	31	32
1	0,4450	0,4407	0,4366	0,4328	0,4291	0,4254	0,4220	0,4188
2	0,3069	0,3043	0,3018	0,2992	0,2968	0,2944	0,2921	0,2898
3	0,2543	0,2533	0,2522	0,2510	0,2499	0,2487	0,2475	0,2463
4	0,2148	0,2151	0,2152	0,2151	0,2150	0,2148	0,2145	0,2141
5	0,1822	0,1836	0,1848	0,1857	0,1864	0,1870	0,1874	0,1878
6	0,1539	0,1563	0,1584	0,1601	0,1616	0,1630	0,1641	0,1651
7	0,1283	0,1316	0,1346	0,1372	0,1395	0,1415	0,1433	0,1449
8	0,1046	0,1089	0,1128	0,1162	0,1192	0,1219	0,1243	0,1265
9	0,0823	0,0876	0,0923	0,0965	0,1002	0,1036	0,1066	0,1093
10	0,0610	0,0672	0,0728	0,0778	0,0822	0,0862	0,0899	0,0931
11	0,0403	0,0476	0,0540	0,0598	0,0650	0,0697	0,0739	0,0777
12	0,0200	0,0284	0,0358	0,0424	0,0483	0,0537	0,0585	0,0629
13	0,0000	0,0094	0,0178	0,0253	0,0320	0,0381	0,0435	0,0485
14	-	-	0,0000	0,0084	0,0159	0,0227	0,0289	0,0344
15	-	-	-	-	0,0000	0,0076	0,0144	0,0206
16	-	-	-	-	-	-	0,0000	0,0068
i	33	34	35	36	37	38	39	
1	0,4156	0,4127	0,4096	0,4068	0,4040	0,4015	0,3989	
2	0,2876	0,2854	0,2834	0,2813	0,2794	0,2774	0,2755	
3	0,2451	0,2439	0,2427	0,2415	0,2403	0,2391	0,2380	
4	0,2137	0,2132	0,2127	0,2121	0,2116	0,2110	0,2104	
5	0,1880	0,1882	0,1883	0,1883	0,1883	0,1881	0,1880	
6	0,1660	0,1667	0,1673	0,1678	0,1683	0,1686	0,1689	
7	0,1463	0,1475	0,1487	0,1496	0,1505	0,1513	0,1520	
8	0,1284	0,1301	0,1317	0,1331	0,1344	0,1356	0,1366	
9	0,1118	0,1140	0,1160	0,1179	0,1196	0,1211	0,1225	
10	0,0961	0,0988	0,1013	0,1036	0,1056	0,1075	0,1092	
11	0,0812	0,0844	0,0873	0,0900	0,0924	0,0947	0,0967	
12	0,0669	0,0706	0,0739	0,0770	0,0798	0,0824	0,0848	
13	0,0530	0,0572	0,0610	0,0645	0,0677	0,0706	0,0733	
14	0,0395	0,0441	0,0484	0,0523	0,0559	0,0592	0,0622	
15	0,0262	0,0314	0,0361	0,0404	0,0444	0,0481	0,0515	
16	0,0131	0,0187	0,0239	0,0287	0,0331	0,0372	0,0409	
17	0,0000	0,0062	0,0119	0,0172	0,0220	0,0264	0,0305	
18	-	-	0,0000	0,0057	0,0110	0,0158	0,0203	
19	-	-	-	-	0,0000	0,0053	0,0101	
20	-	-	-	-	-	-	0,0000	

10. Statistique W du test de normalité de Shapiro-Wilk : coefficients $a_{i,n}$ (suite)

Taille de l'échantillon

i	40	41	42	43	44	45	46
1	0,3964	0,3940	0,3917	0,3894	0,3872	0,3850	0,3830
2	0,2737	0,2719	0,2701	0,2684	0,2667	0,2651	0,2635
3	0,2368	0,2357	0,2345	0,2334	0,2323	0,2313	0,2302
4	0,2098	0,2091	0,2085	0,2078	0,2072	0,2065	0,2058
5	0,1878	0,1876	0,1874	0,1871	0,1868	0,1865	0,1862
6	0,1691	0,1693	0,1694	0,1695	0,1695	0,1695	0,1695
7	0,1526	0,1531	0,1535	0,1539	0,1542	0,1545	0,1548
8	0,1376	0,1384	0,1392	0,1398	0,1405	0,1410	0,1415
9	0,1237	0,1249	0,1259	0,1269	0,1278	0,1286	0,1293
10	0,1108	0,1123	0,1136	0,1149	0,1160	0,1170	0,1180
11	0,0986	0,1004	0,1020	0,1035	0,1049	0,1062	0,1073
12	0,0870	0,0891	0,0909	0,0927	0,0943	0,0959	0,0972
13	0,0759	0,0782	0,0804	0,0824	0,0842	0,0860	0,0876
14	0,0651	0,0677	0,0701	0,0724	0,0745	0,0765	0,0783
15	0,0546	0,0575	0,0602	0,0628	0,0651	0,0673	0,0694
16	0,0444	0,0476	0,0506	0,0534	0,0560	0,0584	0,0607
17	0,0343	0,0379	0,0411	0,0442	0,0471	0,0497	0,0522
18	0,0244	0,0283	0,0318	0,0352	0,0383	0,0412	0,0439
19	0,0146	0,0188	0,0227	0,0263	0,0296	0,0328	0,0357
20	0,0049	0,0094	0,0136	0,0175	0,0211	0,0245	0,0277
21	-	0,0000	0,0045	0,0087	0,0126	0,0163	0,0197
22	-	-	-	0,0000	0,0042	0,0081	0,0118
23	-	-	-	-	-	0,0000	0,0039

Taille de l'échantillon

i	47	48	49	50
1	0,3808	0,3789	0,3770	0,3751
2	0,2620	0,2604	0,2589	0,2574
3	0,2291	0,2281	0,2271	0,2260
4	0,2052	0,2045	0,2038	0,2032
5	0,1859	0,1855	0,1851	0,1847
6	0,1695	0,1693	0,1692	0,1691
7	0,1550	0,1551	0,1553	0,1554
8	0,1420	0,1423	0,1427	0,1430
9	0,1300	0,1306	0,1312	0,1317
10	0,1189	0,1197	0,1205	0,1212
11	0,1085	0,1095	0,1105	0,1113
12	0,0986	0,0998	0,1010	0,1020
13	0,0892	0,0906	0,0919	0,0932
14	0,0801	0,0817	0,0832	0,0846
15	0,0713	0,0731	0,0748	0,0764
16	0,0628	0,0648	0,0667	0,0685
17	0,0546	0,0568	0,0588	0,0608
18	0,0465	0,0489	0,0511	0,0532
19	0,0385	0,0411	0,0436	0,0459
20	0,0307	0,0335	0,0361	0,0386
21	0,0229	0,0259	0,0288	0,0314
22	0,0153	0,0185	0,0215	0,0244
23	0,0076	0,0111	0,0143	0,0174
24	0,0000	0,0037	0,0071	0,0104
25	-	-	0,0000	0,0035

11. Valeurs critiques de la statistique W du test de normalité de Shapiro-Wilk

Quantile					Quantile				
n	1%	2%	5%	10%	n	1%	2%	5%	10%
3	0,753	0,756	0,767	0,789	27	0,894	0,906	0,923	0,935
4	0,687	0,707	0,748	0,792	28	0,896	0,908	0,924	0,936
5	0,686	0,715	0,762	0,806	29	0,898	0,910	0,926	0,937
6	0,713	0,743	0,788	0,826	30	0,900	0,912	0,927	0,939
7	0,730	0,760	0,803	0,838	31	0,902	0,914	0,929	0,940
8	0,749	0,778	0,818	0,851	32	0,904	0,915	0,930	0,941
9	0,764	0,791	0,829	0,859	33	0,906	0,917	0,931	0,942
10	0,781	0,806	0,842	0,869	34	0,908	0,919	0,933	0,943
11	0,792	0,817	0,850	0,876	35	0,910	0,920	0,934	0,944
12	0,805	0,828	0,859	0,883	36	0,912	0,922	0,935	0,945
13	0,814	0,837	0,866	0,889	37	0,914	0,924	0,936	0,946
14	0,825	0,846	0,874	0,895	38	0,916	0,925	0,938	0,947
15	0,835	0,855	0,881	0,901	39	0,917	0,927	0,939	0,948
16	0,844	0,863	0,887	0,906	40	0,919	0,928	0,940	0,949
17	0,851	0,869	0,892	0,910	41	0,920	0,929	0,941	0,950
18	0,858	0,874	0,897	0,914	42	0,922	0,930	0,942	0,951
19	0,863	0,879	0,901	0,917	43	0,923	0,932	0,943	0,951
20	0,868	0,884	0,905	0,920	44	0,924	0,933	0,944	0,952
21	0,873	0,888	0,908	0,923	45	0,926	0,934	0,945	0,953
22	0,878	0,892	0,911	0,926	46	0,927	0,935	0,945	0,953
23	0,881	0,895	0,914	0,928	47	0,928	0,936	0,946	0,954
24	0,884	0,898	0,916	0,930	48	0,929	0,937	0,947	0,954
25	0,888	0,901	0,918	0,931	49	0,929	0,937	0,947	0,955
26	0,891	0,904	0,920	0,933	50	0,930	0,938	0,947	0,955

12. Valeurs critiques de la statistique W' du test de normalité de Shapiro-Francia

$n \backslash \alpha$	0,10	0,05	0,01	$n \backslash \alpha$	0,10	0,05	0,01
5	0,9045	0,8866	0,8336	33	0,9756	0,9678	0,9546
6	0,9119	0,8942	0,8485	34	0,9758	0,9686	0,9548
7	0,9193	0,9018	0,8488	35	0,9760	0,9691	0,9551
8	0,9248	0,9023	0,8676	36	0,9762	0,9695	0,9551
9	0,9302	0,9113	0,8715	37	0,9763	0,9698	0,9556
10	0,9351	0,9186	0,8724	38	0,9765	0,9709	0,9560
11	0,9371	0,9224	0,8889	39	0,9770	0,9712	0,9567
12	0,9409	0,9261	0,8917	40	0,9775	0,9724	0,9576
13	0,9505	0,9288	0,8919	41	0,9779	0,9726	0,9604
14	0,9514	0,9377	0,9063	42	0,9780	0,9730	0,9610
15	0,9523	0,9398	0,9075	43	0,9782	0,9732	0,9611
16	0,9532	0,9400	0,9141	44	0,9785	0,9734	0,9631
17	0,9558	0,9431	0,9143	45	0,9789	0,9743	0,9636
18	0,9561	0,9439	0,9161	46	0,9793	0,9758	0,9643
19	0,9590	0,9493	0,9235	47	0,9801	0,9764	0,9646
20	0,9595	0,9500	0,9283	48	0,9806	0,9767	0,9653
21	0,9617	0,9542	0,9291	49	0,9807	0,9768	0,9655
22	0,9626	0,9544	0,9311	50	0,9810	0,9769	0,9667
23	0,9655	0,9554	0,9329	51	0,9815	0,9772	0,9669
24	0,9662	0,9557	0,9388	52	0,9818	0,9773	0,9673
25	0,9669	0,9572	0,9432	53	0,9822	0,9777	0,9685
26	0,9675	0,9576	0,9437	54	0,9824	0,9780	0,9694
27	0,9688	0,9599	0,9459	55	0,9825	0,9788	0,9695
28	0,9701	0,9637	0,9466	56	0,9826	0,9792	0,9697
29	0,9712	0,9641	0,9490	57	0,9828	0,9796	0,9704
30	0,9723	0,9651	0,9506	58	0,9830	0,9797	0,9705
31	0,9734	0,9654	0,9524	59	0,9833	0,9800	0,9707
32	0,9745	0,9669	0,9538	60	0,9836	0,9801	0,9710

12. Valeurs critiques de la statistique W' du test de normalité de Shapiro-Francia (suite)

$n \backslash \alpha$	0,10	0,05	0,01
61	0,9838	0,9810	0,9718
62	0,9839	0,9811	0,9722
63	0,9842	0,9812	0,9722
64	0,9843	0,9813	0,9728
65	0,9844	0,9813	0,9732
66	0,9849	0,9817	0,9734
67	0,9852	0,9820	0,9736
68	0,9852	0,9821	0,9739
69	0,9854	0,9821	0,9744
70	0,9855	0,9822	0,9755
71	0,9856	0,9827	0,9757
72	0,9857	0,9828	0,9759
73	0,9859	0,9830	0,9762
74	0,9862	0,9832	0,9764
75	0,9864	0,9834	0,9765
76	0,9866	0,9834	0,9771
77	0,9867	0,9837	0,9774
78	0,9868	0,9839	0,9776
79	0,9869	0,9841	0,9779
80	0,9871	0,9841	0,9780
81	0,9872	0,9845	0,9782
82	0,9873	0,9847	0,9785
83	0,9874	0,9848	0,9786
84	0,9876	0,9848	0,9787
85	0,9877	0,9849	0,9788
86	0,9877	0,9851	0,9790
87	0,9878	0,9853	0,9791
88	0,9880	0,9854	0,9792
89	0,9880	0,9855	0,9793
90	0,9882	0,9857	0,9794
91	0,9883	0,9859	0,9796
92	0,9887	0,9859	0,9800
93	0,9888	0,9861	0,9802
94	0,9888	0,9862	0,9805
95	0,9889	0,9862	0,9807
96	0,9889	0,9863	0,9811
97	0,9890	0,9867	0,9812
98	0,9893	0,9868	0,9814
99	0,9894	0,9872	0,9818
100	0,9896	0,9873	0,9820
101	0,9896	0,9873	0,9821
102	0,9897	0,9873	0,9821
103	0,9897	0,9875	0,9822
104	0,9897	0,9876	0,9822
105	0,9898	0,9876	0,9823
106	0,9899	0,9877	0,9823
107	0,9899	0,9878	0,9825
108	0,9899	0,9878	0,9826
109	0,9900	0,9878	0,9828
110	0,9901	0,9879	0,9832
111	0,9902	0,9880	0,9834
112	0,9902	0,9882	0,9835
113	0,9906	0,9885	0,9837
114	0,9906	0,9886	0,9837
115	0,9907	0,9887	0,9839
116	0,9909	0,9887	0,9840

$n \backslash \alpha$	0,10	0,05	0,01
117	0,9909	0,9887	0,9840
118	0,9910	0,9888	0,9842
119	0,9911	0,9888	0,9843
120	0,9913	0,9890	0,9844
125	0,9915	0,9894	0,9853
130	0,9918	0,9897	0,9859
135	0,9920	0,9900	0,9863
140	0,9922	0,9904	0,9865
145	0,9924	0,9906	0,9871
150	0,9925	0,9908	0,9875
155	0,9927	0,9911	0,9880
160	0,9931	0,9919	0,9887
165	0,9932	0,9920	0,9890
170	0,9934	0,9921	0,9892
175	0,9935	0,9922	0,9896
180	0,9936	0,9923	0,9899
185	0,9937	0,9925	0,9902
190	0,9939	0,9926	0,9903
195	0,9940	0,9929	0,9904
200	0,9941	0,9931	0,9908
210	0,9945	0,9934	0,9910
220	0,9946	0,9935	0,9913
230	0,9948	0,9937	0,9917
240	0,9951	0,9940	0,9920
250	0,9953	0,9944	0,9923
260	0,9955	0,9946	0,9924
270	0,9956	0,9947	0,9928
280	0,9957	0,9949	0,9930
290	0,9959	0,9950	0,9932
300	0,9959	0,9951	0,9934
310	0,9960	0,9952	0,9935
320	0,9962	0,9955	0,9937
330	0,9964	0,9956	0,9940
340	0,9964	0,9957	0,9941
350	0,9965	0,9958	0,9943
360	0,9965	0,9958	0,9943
370	0,9967	0,9960	0,9945
380	0,9968	0,9961	0,9946
390	0,9969	0,9961	0,9947
400	0,9969	0,9962	0,9951
420	0,9970	0,9963	0,9953
440	0,9972	0,9967	0,9955
460	0,9973	0,9967	0,9958
480	0,9974	0,9969	0,9959
500	0,9975	0,9971	0,9959
550	0,9978	0,9974	0,9961
600	0,9979	0,9975	0,9963
650	0,9979	0,9978	0,9967
700	0,9981	0,9979	0,9972
750	0,9983	0,9979	0,9973
800	0,9984	0,9980	0,9974
850	0,9985	0,9981	0,9975
900	0,9986	0,9983	0,9976
950	0,9987	0,9984	0,9978
999	0,9988	0,9985	0,9979
1000	0,9988	0,9985	0,9979

Dans la vie quotidienne, il est bien rare qu'un sondage ne vienne pas troubler nos certitudes ! Ils sont partout : journaux, ondes FM, internet et télévision !

La technique des sondages tient à la fois de la sociologie et de la statistique. De plus, elle a un vocabulaire spécifique. D'ailleurs, pour être certain que nous parlons bien tous de la même chose, précisons le mot « sondage ». En France, le mot sondage désigne à la fois l'enquête par sondage et le sondage d'opinion. Le premier s'appuie sur une théorie probabiliste, le second est purement empirique. Il faut distinguer les sondages aléatoires des sondages non aléatoires ou encore parfois appelés « méthodes empiriques » ou « méthodes par choix raisonné ». Dans ces dernières, nous trouvons la fameuse méthode des quotas, largement utilisée par les instituts privés qui nous donnent quotidiennement la tendance de l'opinion. Mais ce n'est pas la seule méthode empirique. Nous citerons : la méthode des itinéraires, la méthode des unités-types. Nous renvoyons le lecteur à l'ouvrage de Pascal Ardilly *Les Techniques de sondage*.

Le mot « sondage » n'est pas le seul qui prête à confusion, il en va de même des mots « hasard » et le mot « aléatoire ». Dans le langage courant, un résultat est « aléatoire » ou un événement arrive « par hasard » lorsqu'il est imprévu, inattendu, subi. En probabilités et en statistique, les variables aléatoires et leur loi de probabilité sont des objets bien identifiés et le « hasard » est construit selon une loi de probabilité. En théorie des sondages, tirer un échantillon aléatoire ou probabiliste c'est extraire un échantillon de la population selon une loi de probabilité, sur l'ensemble des échantillons, que nous nous sommes fixée à l'avance (cette phase est appelée « plan de sondage »). En particulier, tirer un échantillon aléatoire simple à probabilités égales de taille n (nous disons abusivement au lycée « tirer au hasard ») signifie que nous effectuons cette opération de telle façon que tous les échantillons de taille n aient la même probabilité d'être tirés. Ceci n'est possible qu'en suivant des procédures aléatoires très contrôlées, par exemple, en reportant dans une liste les identifiants de chacun des individus de la population (liste appelée « base de sondage ») et en utilisant de façon adéquate des tables de nombres aléatoires ou le générateur de nombres « pseudo-aléatoires » de sa calculatrice pour le tirage des individus.

Mais pourquoi sonder ?

Parfois, certaines situations ne permettent pas de faire un recensement. Prenons l'exemple du recensement des êtres marins dans un lac. Comment les experts peuvent-ils connaître le nombre exact d'individus qui vivent dans ce milieu ? Il faudrait assécher le lac mais alors tous les poissons mourraient. Bien entendu, des méthodes ont été inventées pour donner l'estimation du nombre de vies dans ce milieu, comme la méthode de capture-recapture¹. Il faut aussi avoir conscience qu'un recensement coûte cher en termes d'argent et de temps. Le premier recensement date de Jules César. Il est aussi intéressant de savoir que, jusqu'en 2004, le « dénombrement » de la population française se faisait par « recensement » tous les sept à neuf ans ;

¹ Voir l'Express de Statistique de D. Fredon, M. Maumy-Bertrand et F. Bertrand, aux éditions Dunod.

il se fait depuis par « sondage aléatoire » et par rotation tous les ans ; ce nouveau dénombrement permet d'obtenir depuis 2008 une amélioration sensible de la qualité de l'information.

Marge d'erreur

Lorsque les journalistes publient les chiffres, ils ne nous donnent pas l'intervalle de confiance mais l'estimation ponctuelle. Or comme nous pouvons le voir dans la Fiche 106, donner une estimation ponctuelle est insuffisant. Qu'appelons-nous la marge d'erreur ? En général, cette marge d'erreur est reliée avec la notion fondamentale d'estimation par intervalle de confiance à 95 % d'un paramètre. Dans les sondages, le paramètre est souvent la proportion. Comme mentionné dans la Fiche 110, nous pouvons établir un intervalle de confiance pour une proportion. Si nous prenons la formule la plus simple, en supposant que les approximations faites sont réalisables à savoir l'approximation de la loi binomiale par la loi normale est réaliste, Fiche 94, et que le taux de sondage défini par n/N est négligeable (c'est-à-dire assimiler le tirage sans remise à un tirage avec remise), nous pouvons alors dire que la marge d'erreur à 95 % de confiance est égale à $2 \sqrt{p(1-p)/n}$. En fait, il faudrait plutôt parler de marge d'incertitude. Dans la vie quotidienne, cette marge d'erreur est exprimée en « points » et non en %. Si nous observons une proportion égale à 52 % sur un échantillon de taille 1 000 avec une marge d'incertitude de 3 %, au lieu de laisser croire que la vraie proportion inconnue est quasiment égale à 52 % il faudrait annoncer que « la vraie proportion est comprise entre 49 % et 55 % » et annoncer de plus que cette affirmation n'est pas certaine, que cet intervalle a été construit avec un niveau de confiance de 95 %. Nous pouvons donc nous tromper dans 5 % des cas, c'est-à-dire dans 5 % des cas, la vraie valeur inconnue de la proportion n'appartient pas à l'intervalle construit. Dernier point essentiel : il faut remarquer que la précision de l'estimation du paramètre ne dépend pas de la taille de la population. Pour savoir si la soupe est salée, il suffit de goûter une cuillerée, que cette cuillère soit extraite d'un petit bol ou d'un très grand chaudron.

Conclusion

Si cette théorie des sondages est parfaitement connue et maîtrisée, pourquoi les sondages publiés dans les journaux ou dans tout autre type de média se trompent-ils ? L'exemple le plus frappant de ces dernières années est le « surprenant » résultat du premier tour de l'élection présidentielle du 21 avril 2002 : le second tour opposa Jacques Chirac à Jean-Marie Le Pen et non à Lionel Jospin comme l'annonçaient les sondeurs. Se seraient-ils encore trompés ? Michel Lejeune répond à cette attaque dans l'article du *Monde* du 26 avril 2002 : « Il faut savoir qu'il n'y a aucun statisticien spécialiste de la théorie des sondages dans les cinq plus grands instituts français. Aujourd'hui la communauté scientifique doit se sentir en partie responsable de ce qui vient de se produire. Par dédain, elle n'a jamais voulu s'intéresser à la pratique des sondages, ni vraiment d'ailleurs, à la théorie. » Effectivement, pendant des années, la pratique des sondages était considérée par les statisticiens universitaires comme une activité ne concernant que la statistique publique (INSEE, INED...), ne présentant donc aucun intérêt pour la formation des étudiants, ni sur un plan pédagogique, ni pour une recherche théorique.

Corrigés des exercices

Les corrigés des exercices signalés dans les fiches par  sont disponibles sur le site dunod.com à partir de la page de présentation de l'ouvrage.

Fiche 1

1.1 La négation de $x = 2$ est $x \neq 2$.

1.2 Les contraposées sont :

1. $x = 0$ ou $y = 0 \Rightarrow xy = 0$.

2. $n \neq 2$ et n pair $\Rightarrow n$ n'est pas un nombre premier.

1.3 n est pair équivaut à $n = 2k$ avec $k \in \mathbb{Z}$. Comme $n^2 = 4k^2 = 2 \times 2k^2$, on obtient que n^2 est également pair.

1.4 Les valeurs de vérité sont : F, V, F et F .

1.5 Les valeurs de vérité sont : V, V, V et F .

Fiche 2

2.1 La contraposée de l'assertion que nous souhaitons montrer est : si n est pair alors n^2 est pair. Supposons donc que n soit un nombre pair. Alors $n = 2k$ avec $k \in \mathbb{Z}$. Nous en déduisons que $n^2 = 4k^2 = 2 \times 2k^2$ donc nous montrons ici que n^2 est forcément un nombre pair.

2.2 La contraposée de l'assertion que nous souhaitons montrer est : si $\frac{1}{x-1} = \frac{1}{y-1}$ alors $x = y$.

Supposons donc que $\frac{1}{x-1} = \frac{1}{y-1}$, en multipliant de chaque côté de l'égalité par $(x-1)(y-1)$ nous avons : $y-1 = x-1$ et du coup $x = y$.

2.3 Pour $n = 5$, nous avons $2^{2^5} + 1 = 4294967297$, qui est divisible par 641 et qui donc n'est pas un nombre premier. Ce contre-

exemple pour $n = 5$ montre donc que $2^{2^n} + 1$ n'est pas un nombre premier pour tout $n \in \mathbb{N}^*$.

2.5 Tout d'abord, pour $n = 0$, $2^n = 1 \geq 0 + 1 = 1$. Supposons que l'inégalité soit vraie à un ordre fixé $n \in \mathbb{N}$. Montrons qu'elle reste vraie à l'ordre $n+1$. $2^{n+1} = 2 \times 2^n \geq 2 \times (n+1)$ par hypothèse de récurrence. Nous en déduisons que $2^{n+1} \geq 2n+2 \geq n+2$. La propriété étant vraie à l'ordre $n+1$ avec $n \in \mathbb{N}$, elle est vraie pour tout $n \in \mathbb{N}$.

2.6 Tout d'abord, pour $n = 4$, $2^n = 16 \leq 4! = 24$. Supposons que l'inégalité soit vraie à un ordre fixé $n \in \mathbb{N}$, $n \geq 4$. Montrons qu'elle reste vraie à l'ordre $n+1$. $2^{n+1} = 2 \times 2^n \leq 2 \times n!$ par hypothèse de récurrence. Comme $n \geq 4$, $2 \leq n+1$ et donc $2^{n+1} \leq (n+1)n! = (n+1)!$. La propriété étant vraie à l'ordre $n+1$ avec $n \in \mathbb{N}$, $n \geq 4$, elle est vraie pour tout $n \in \mathbb{N}$, $n \geq 4$.

2.7 Tout d'abord, pour $n = 1$, $3^{2^{n+2}} - 2^{n+1} = 77$ qui est bien divisible par 7. Supposons que la propriété soit vraie à un ordre fixé $n \in \mathbb{N}$, $n \geq 1$. Montrons qu'elle reste vraie à l'ordre $n+1$. $3^{2^{(n+1)+2}} - 2^{(n+1)+1} = 3^2 \times 3^{2^{n+2}} - 2 \times 2^{n+1} = 9 \times 3^{2^{n+2}} - 2 \times 2^{n+1} = 2 \times (3^{2^{n+2}} - 2^{n+1}) + 7 \times 3^{2^{n+2}}$. Par hypothèse de récurrence, $2 \times (3^{2^{n+2}} - 2^{n+1})$ est divisible par 7. D'autre part, $7 \times 3^{2^{n+2}}$ est divisible par 7 donc $3^{2^{(n+1)+2}} - 2^{(n+1)+1}$ est divisible par 7 et la propriété est vraie à l'ordre $n+1$. La propriété étant vraie à l'ordre $n+1$ avec $n \in \mathbb{N}$, $n \geq 1$, elle est vraie pour tout $n \in \mathbb{N}$, $n \geq 1$.

2.8 Tout d'abord, pour $n = 1$, $n(2n+1)(7n+1) = 24$ qui est bien divisible par 6. Supposons que la propriété soit vraie à un ordre fixé $n \in \mathbb{N}$, $n \geq 1$. Montrons qu'elle reste vraie à l'ordre $n+1$. $(n+1)(2(n+1)+1)(7(n+1)+1) = (n+1)(2n+3)(7n+8) = 14n^3 + 51n^2 + 61n + 24$.

D'autre part, $n(2n+1)(7n+1) = 14n^3 + 9n^2 + n$. Nous en déduisons que $(n+1)(2(n+1)+1)(7(n+1)+1) = n(2n+1)(7n+1) + 42n^2 + 60n + 24$. Par hypothèse de récurrence, $n(2n+1)(7n+1)$ est divisible par 6. D'autre part, $42n^2 + 60n + 24 = 6 \times (7n^2 + 10n + 4)$ est bien divisible par 6 donc $(n+1)(2(n+1)+1)(7(n+1)+1)$ est divisible par 6 et la propriété est vraie à l'ordre $n+1$. La propriété étant vraie à l'ordre $n+1$ avec $n \in \mathbb{N}, n \geq 1$, elle est vraie pour tout $n \in \mathbb{N}, n \geq 1$.

2.9 Pour $n = 1$, $\sum_{k=1}^1 \frac{1}{k(k+1)} = \frac{1}{2}$ et $\frac{n}{n+1} = \frac{1}{2}$ donc nous avons bien l'égalité. Supposons que l'égalité soit vraie à l'ordre n et montrons qu'elle reste vraie à l'ordre $n+1$. $\sum_{k=1}^{n+1} \frac{1}{k(k+1)} = \sum_{k=1}^n \frac{1}{k(k+1)} + \frac{1}{(n+1)(n+2)} = \frac{n}{n+1} + \frac{1}{(n+1)(n+2)} = \frac{(n+1)^2}{(n+1)(n+2)} = \frac{(n+1)}{(n+2)}$. L'égalité est vraie à l'ordre $n+1$, elle est donc vraie $\forall n \in \mathbb{N}^*$.

2.10 Pour $n = 1$, $\sum_{k=1}^1 (2k-1) = 1$ et $n^2 = 1$ donc nous avons bien l'égalité. Supposons que l'égalité soit vraie à l'ordre n et montrons qu'elle reste vraie à l'ordre $n+1$. $\sum_{k=1}^{n+1} (2k-1) = \sum_{k=1}^n (2k-1) + 2(n+1) - 1 = n^2 + 2n + 1 = (n+1)^2$. L'égalité est vraie à l'ordre $n+1$, elle est donc vraie $\forall n \in \mathbb{N}^*$.

Fiche 3

3.1 $A \cap B = \{2; 3\}$ et $A \cup B = \{1; 2; 3; 4\}$.

3.2 $\bar{A} = \{x \in \mathbb{R} : x \leq 9\}$.

3.3 Nous pouvons choisir $A = \{1; 2; 3; 4\}$, $B = \{1\}$ et $C = \{2; 3; 4\}$. En effet, $A \cup B = A \cup C = \{1; 2; 3; 4\}$ et $B \neq C$.

3.4 Montrons d'abord que $A \subset B$. Soit $x \in A$, alors $x \in A \cup B$ et comme $A \cup B = A \cap B$, nous avons forcément $x \in B$.

Montrons que $B \subset A$. Soit $x \in B$, alors $x \in A \cup B$ et comme $A \cup B = A \cap B$, nous avons forcément $x \in A$.

La double inclusion nous permet de conclure que $A = B$.

3.5 Soit $x \in B$, alors $x \in A \cup B$ et comme $A \cup B = A \cup C$, $x \in A \cup C$. Nous avons alors deux possibilités. Si $x \in A$, comme $x \in B$, $x \in A \cap B$ et comme $A \cap B = A \cap C$, nous en déduisons que $x \in C$. Donc $B \subset C$. Sinon $x \in C$ et nous avons directement le résultat cherché.

Montrons maintenant que $C \subset B$. Soit $x \in C$, alors $x \in A \cup C$ et comme $A \cup B = A \cup C$, $x \in A \cup B$. Nous avons alors deux possibilités. Si $x \in A$, comme $x \in C$, $x \in A \cap C$ et comme $A \cap B = A \cap C$, nous en déduisons que $x \in B$. Donc $C \subset B$. Sinon $x \in B$ et nous avons directement le résultat cherché.

La double inclusion permet de conclure que $B = C$.

3.6

1. $A \cap B$. 3. $\bar{A} \cup B$.

2. $A \cup B$. 4. $\overline{A \cup B}$.

Fiche 4

4.1 La relation n'est pas réflexive car x n'est pas en relation avec lui-même puisque $x = x$ et non $x \neq x$. Cela suffit à montrer que cette relation n'est pas une relation d'ordre.

4.2

1. La relation n'est pas réflexive. Elle n'est pas transitive. Ce n'est donc ni une relation d'équivalence ni une relation d'ordre.

2. La relation est réflexive. Elle est également symétrique et transitive. La relation est donc une relation d'équivalence. Ce n'est pas en revanche une relation d'ordre car elle n'est pas anti-symétrique.

3. La relation est réflexive. Elle n'est pas symétrique. Elle est également transitive. La relation n'est donc pas une relation d'équivalence. En revanche, elle est anti-symétrique. C'est donc une relation d'ordre.

4.3 La relation est réflexive car $x^2 - x^2 = 0 = x - x$. Elle est symétrique puisque si $x^2 - y^2 = x - y$ nous avons $y^2 - x^2 = y - x$. Elle est enfin transitive car si $x \mathcal{R} y$ et $y \mathcal{R} z$, alors $x^2 - z^2 = x^2 - y^2 + y^2 - z^2 = x - y + y - z = x - z$ donc $x \mathcal{R} z$. C'est donc une relation d'équivalence. Pour trouver la classe d'équivalence nous devons résoudre $x^2 - y^2 = x - y \Rightarrow (x - y)(x + y - 1) = 0$. L'ensemble des y qui vérifie cette équation est de la forme $\{x; 1 - x\}$. La classe d'équivalence est donc composée de deux éléments x et $1 - x$ sauf dans le cas où $x = \frac{1}{2}$ pour lequel $x = 1 - x$.

4.4

1. Cette partie est non vide minorée mais non majorée. Sa borne inférieure est a .

2. Cette partie est non vide minorée et majorée. Sa borne inférieure est a et sa borne supérieure est $a + b$.

3. Cet ensemble ne prend que deux valeurs quelles que soient les valeurs de n : $a - b$ et $a + b$. La borne inférieure est $a - b$ et la borne supérieure $a + b$.

4.5 Il existe b tel que $\forall a \in A, a \leq b$. Donc A est une partie non vide et majorée et $\sup(A) \leq b$. D'autre part, B est non vide et minorée donc $\sup(A) \leq \inf(B)$.

Fiche 5

5.1 $f_3 \circ f_1 = x^2 - 4, f_1 \circ f_3 = (x - 4)^2, f_3 \circ f_2 = \sqrt{x} - 4, f_2 \circ f_3 = \sqrt{x - 4}$ et $f_1 \circ f_1 = x^4$.

5.2 $f \circ g = 3(x^2 - 1) + 1 = 3x^2 - 2$ et $g \circ f = (3x + 1)^2 - 1 = 9x^2 + 6x$ donc $f \circ g \neq g \circ f$.

5.3 L'application est à la fois injective et surjective, elle est donc bijective.

5.4 f est injective car si $f(n) = f(n')$ alors $n + 1 = n' + 1$ et $n = n'$. Par contre, f n'est pas surjective car 0 n'a pas d'antécédent puisque si on résout $0 = n + 1$ alors $n = -1$ et $-1 \notin \mathbb{N}$. L'application ne sera donc pas bijective.

5.5 L'application f est injective car si $f(x; y) = f(x'; y')$ alors $x + y = x' + y'$ et $x - y = x' - y'$. Si on ajoute ces deux dernières équations, nous obtenons $2x = 2x'$ et donc $x = x'$ et au contraire, si on soustrait ces deux équations, nous obtenons directement $2y = 2y'$ et donc $y = y'$. L'application est également surjective car pour $(x'; y') \in \mathbb{R}^2$ tel que $f(x; y) = (x'; y')$, il existe $x = \frac{x' + y'}{2}$ et $y = \frac{x' - y'}{2}$ comme antécédents qui sont bien dans $\mathbb{R}^2$. Nous en déduisons donc que cette application est bien bijective.

5.6

1. Supposons que $\forall a, a' \in A, f(a) = f(a')$. Nous voulons montrer que $a = a'$. Comme $f(a) = f(a'), g(f(a)) = g(f(a')) \Rightarrow g \circ f(a) = g \circ f(a')$. Comme $g \circ f$ est supposée injective, nous avons directement $a = a'$.

2. Soit $c \in C$. Nous voulons montrer qu'il existe $b \in B$ tel que $g(b) = c$. Comme $g \circ f$ est surjective, pour $c \in C$ il existe $a \in A$ tel que $c = g \circ f(a)$. Il nous suffit de poser $b = f(a)$ et nous obtenons bien $c = g(b)$ donc tout élément $c \in C$ possède bien un antécédent $b \in B$ tel que $g(b) = c$ et donc g est surjective.

5.7 La fonction tangente est strictement croissante et continue sur $] -\frac{\pi}{2}; \frac{\pi}{2}[$. De plus, lorsque x varie sur $] -\frac{\pi}{2}; \frac{\pi}{2}[$, $\tan(x)$ varie de $-\infty$ à $+\infty$. Elle réalise donc une bijection de $] -\frac{\pi}{2}; \frac{\pi}{2}[$ sur $] -\infty; +\infty[$. Elle admet donc une bijection réciproque $\tan^{-1}$ telle que pour $y \in] -\infty; +\infty[$, il existe $x \in] -\frac{\pi}{2}; \frac{\pi}{2}[$ tel que $y = \tan x$. $\tan^{-1}$ est souvent notée Arctan .

Fiche 6

6.1 Nous utilisons le raisonnement par récurrence. Pour $n = 4$, nous avons $2^n = 2^4 = 16$ et $n! = 4! = 24$ donc nous avons bien $2^2 < n!$. Supposons que l'inégalité soit vraie à l'ordre n et essayons de la montrer à l'ordre $n + 1$. Nous avons, $2^{n+1} = 2 \times 2^n < 2 \times n! < (n + 1) \times n! = (n + 1)!$. L'inégalité étant vraie à l'ordre $n + 1$, elle est vraie $\forall n \geq 4$.

6.2 Nous montrons cette affirmation par récurrence. Tout d'abord, pour $n = 1$, $3 \times 5^1 + 2^1 = 17$ est bien divisible par 17. Supposons que cette assertion soit vraie à l'ordre n et montrons qu'elle reste vraie à l'ordre $n + 1$. $3 \times 5^{2(n+1)-1} + 2^{3(n+1)-2} = 8 \times (3 \times 5^{2n-1} + 2^{3n-1}) + 17 \times 3 \times 5^{2n-1}$. Par hypothèse de récurrence, la première partie de ce dernier terme est divisible par 17. La deuxième partie, soit $17 \times 3 \times 5^{2n-1}$ est également divisible par 17. Nous en déduisons que la propriété est vraie à l'ordre $n + 1$, elle est donc vraie $\forall n \in \mathbb{N}^*$.

6.3 Pour $n = 1$, $1 = \frac{1(1+1)}{2}$ donc la propriété est vraie. Supposons qu'elle soit vraie à l'ordre n et montrons qu'elle reste vraie à l'ordre $n + 1$. Nous avons $1 + 2 + \dots + n + 1 = \frac{n(n+1)}{2} + n + 1 = \frac{n(n+1) + 2(n+1)}{2} = \frac{(n+1)(n+2)}{2}$ donc la propriété est également vraie à l'ordre $n + 1$, elle est vraie pour tout $n \in \mathbb{N}^*$.

6.4 Pour $n = 1$, $1^2 = 1 = \frac{1(1+1)(2 \times 1 + 1)}{6}$. Supposons que la propriété soit vraie à l'ordre n et montrons qu'elle reste vraie à l'ordre $n + 1$.

$1^2 + 2^2 + \dots + (n+1)^2 = 1^2 + 2^2 + \dots + n^2 + (n+1)^2 = \frac{n(n+1)(2n+1)}{6} + (n+1)^2 = \frac{n(n+1)(2n+1) + 6(n+1)^2}{6} = \frac{(n+1)(n(2n+1) + 6(n+1))}{6} = \frac{(n+1)(n+2)(2n+3)}{6}$. Nous en déduisons donc que l'égalité est vraie à l'ordre $n+1$ et donc $\forall n \in \mathbb{N}^*$.

6.5 D'une manière générale, les nombres proposés sont de la forme $n! + k$ avec $2 \leq k \leq n$. Ils peuvent tous s'écrire $n! + k = k \times m + k$ avec $m \in \mathbb{N}^*$. Ils sont donc divisibles par k et ne sont donc pas des nombres premiers.

6.6 $46848 = 2^8 \times 3 \times 61$.

6.7 Le cardinal de E est 4 car cet ensemble est composé de 4 éléments.

6.8 Il suffit de choisir une fonction constante sur un ensemble A infini.

Fiche 7

7.1

1. Nous avons $\forall a, b, c \in \mathbb{R}, (a * b) * c = (a + b + ab) * c = a + b + ab + c + ac + bc + abc = a * (b + c + bc) = a * (b * c)$. Nous en déduisons que $*$ est associative.

2. Nous avons $\forall a, b \in \mathbb{R}, a * b = a + b + ab = b + a + ba = b * a$. Nous en déduisons que $*$ est commutative.

3. Nous avons $0 * b = b * 0 = b$ donc $*$ possède un élément neutre qui est 0.

7.2 Nous avons tout d'abord $\forall x \in G, x * 0 = x$ et $x * (-x) = 0$ donc 0 est neutre et tout élément de G est inversible. D'autre part, nous avons l'associativité puisque $\forall x, y, z \in G, x * (y * z) = x + y + z + \frac{xyz}{xy + xz + yz + 1} = (x * y) * z$. Nous en déduisons que $(G, *)$ est un groupe.

7.3 Non. Un élément $x \in \mathbb{Z}^*$ n'a pas de symétrique.

7.4 En effet, tout élément non nul de $\mathbb{N}$ n'a pas de symétrique dans $\mathbb{N}$.

7.6

1. Nous avons :

- $Z(F)$ est un sous-ensemble de F .
- $1_F \in Z(F)$ car $\forall y \in F, 1_F * y = y = y * 1_F$.
- $\forall x, x' \in Z(F), \forall y \in F, (x * x') * y = x * x' * y = x * y * x' = y * x * x' = y * (x * x')$.

- $\forall x \in Z(F), \forall y \in F, xy = yx$ donc $x^{-1}xy = x^{-1}yx, y = x^{-1}yx$ et $yx^{-1} = x^{-1}y$. Nous en déduisons que $x^{-1} \in Z(F)$.

$Z(F)$ est donc un sous-groupe de F .

2. F est commutatif équivaut à dire que $\forall x \in F, \forall y \in F, x * y = y * x$ ce qui revient à ce que $Z(F) = F$.

Fiche 8

8.1

1. Puisque $x^n = 0$, nous avons $1 = 1 - x^n = (1 - x) \times (1 + x + \dots + x^{n-1}) = (1 + x + \dots + x^{n-1}) \times (1 - x)$. Par suite, $1 - x$ est inversible et son inverse est $1 + x + \dots + x^{n-1}$.

2. Si $x^n = 0$, nous avons $(xy)^n = x^n \times y^n = 0$. Nous en déduisons que xy est nilpotent.

8.2

Supposons que x soit inversible à droite. Alors, il existe $y \in A$ tel que $x \times y = 1$. Si nous multiplions respectivement cette dernière équation à gauche et à droite par y et x , nous obtenons $y \times x \times y \times x = y \times x$. Cela nous permet d'écrire que $y \times x \times (y \times x - 1) = 0$. Enfin, puisque A est intègre et x et y non nuls, nous en déduisons que $y \times x - 1 = 0$ et donc $y \times x = 1$. x est donc inversible à gauche.

8.3

Tout d'abord, $H \subset A$. D'autre part, $1_A \in H$ puisque $\forall y \in A, 1_A y = y 1_A$. $\forall x, y \in H, (x + y)a = xa + ya = ax + ay = a(x + y)$ donc $x + y \in H$. Enfin, $(x \times y)a = x \times y \times a = x \times a \times y = a \times x \times y$ donc $x \times y \in H$ et $(-x)a = -x \times a = a \times -x$ donc $-x \in H$. Nous en déduisons que H est un sous-anneau de A .

8.4

Soient $(A, +_A, \times_A), (B, +_B, \times_B)$ et $(C, +_C, \times_C)$ des anneaux et $f : A \rightarrow B$ et $g : B \rightarrow C$ deux morphismes d'anneaux. Alors,

- $(g \circ f)(1_A) = g(f(1_A)) = g(1_B) = 1_C$.
- $\forall x, y \in A, (g \circ f)(x +_A y) = g(f(x +_A y)) = g(f(x) +_B f(y)) = g(f(x)) +_C g(f(y)) = (g \circ f)(x) +_C (g \circ f)(y)$.
- $\forall x, y \in A, (g \circ f)(x \times_A y) = g(f(x \times_A y)) = g(f(x) \times_B f(y)) = g(f(x)) \times_C g(f(y)) = (g \circ f)(x) \times_C (g \circ f)(y)$.

Nous en déduisons que $g \circ f$ est un morphisme d'anneaux.

8.5 $(\mathbb{Z}, +, \times)$ n'est pas un corps car seuls les éléments 1 et -1 sont inversibles pour la multiplication.

Fiche 9

9.1

1. 1. 2. 2. 3. 31.

9.2

$a = 3, b = 14$ et $c = 12$.

9.3

$6n + 5$.

9.4

Si x est pair, il existe alors $p \in \mathbb{Z}$ tel que $x = 2p$. Donc $x^2 = 4p^2$. Nous avons alors deux cas de figure. Le premier cas est si l'on considère que p est pair. Alors, p^2 est pair et donc x^2 est divisible par 8 c'est-à-dire $x^2 \equiv 0[8]$. Si par contre p est impair, p^2 est impair et donc x^2 est divisible par 4 c'est-à-dire $x^2 \equiv 4[8]$.

9.5

La solution générale est donc le couple $(5 - 23k; -8 + 37k)$ avec $k \in \mathbb{Z}$.

Fiche 10

10.1

1. Faux. En effet, ce produit scalaire peut également valoir $-AB \times AC$ donc les points A, B et C sont alignés si et seulement si : $|\overrightarrow{AB} \cdot \overrightarrow{AC}| = AB \times AC$.

2. Vrai.

3. Vrai.

10.2

Nous devons montrer que le produit scalaire entre $(\vec{u} + \vec{v})$ et $(\vec{u} - \vec{v})$ est nul. Nous avons, $(\vec{u} + \vec{v}) \cdot (\vec{u} - \vec{v}) = \|\vec{u}\|^2 - \|\vec{v}\|^2 = 0$ car les vecteurs $\vec{u}$ et $\vec{v}$ sont supposés avoir la même norme. Les vecteurs $(\vec{u} + \vec{v})$ et $(\vec{u} - \vec{v})$ sont donc orthogonaux.

10.3

Nous avons $(\vec{u} \wedge \vec{v}) \wedge \vec{w} = -\vec{i}$ et $\vec{u} \wedge (\vec{v} \wedge \vec{w}) = -\vec{j}$. Ce résultat permet de montrer que le produit vectoriel n'est pas en général associatif.

10.4

Pour cela, nous avons besoin de la formule de Gibbs dont nous donnons l'expression. Soient $\vec{u}, \vec{v}$ et $\vec{w}$ trois vecteurs de $\mathbb{R}^3$. Alors, nous avons,

$$\vec{u} \wedge (\vec{v} \wedge \vec{w}) = (\vec{u} \cdot \vec{w}) \cdot \vec{v} - (\vec{u} \cdot \vec{v}) \cdot \vec{w}.$$

Si nous utilisons cette formule, nous avons :
 $\vec{u} \wedge (\vec{v} \wedge \vec{w}) = (\vec{u} \cdot \vec{w}) \cdot \vec{v} - (\vec{u} \cdot \vec{v}) \cdot \vec{w}.$
 $\vec{v} \wedge (\vec{w} \wedge \vec{u}) = (\vec{v} \cdot \vec{u}) \cdot \vec{w} - (\vec{v} \cdot \vec{w}) \cdot \vec{u}.$
 $\vec{w} \wedge (\vec{u} \wedge \vec{v}) = (\vec{w} \cdot \vec{v}) \cdot \vec{u} - (\vec{w} \cdot \vec{u}) \cdot \vec{v}.$

En sommant ces trois équations, comme le produit scalaire est commutatif, nous trouvons le vecteur nul.

10.5

1. $(\vec{u}, \vec{v}, \vec{w}) = 70$.

2. $(\vec{u}, \vec{w}, \vec{v}) = -70$.

3. $(\vec{v}, \vec{w}, \vec{u}) = 70$.

4. $(\vec{u}, \vec{v}, \vec{u}) = 0$.

Fiche 11

11.1

L'angle aigu α est tel que $\cos(\alpha) = \frac{1}{2}$. Nous en déduisons que $\alpha = \frac{\pi}{3}$.

11.2

L'équation du plan passant par ces trois points est : $5x - 5y - 5 = 0$.

11.3

L'équation du plan passant par le point A et dont le vecteur directeur est $\vec{n}$ est : $x + y + 2z + 3 = 0$.

11.4

Ils sont sécants.

11.5

1. L'équation de la sphère est : $(x - \frac{1}{2})^2 + (y - 1)^2 + (z + 2)^2 = 3^2$.

2. Le centre de la sphère est le point de coordonnées $\begin{pmatrix} \frac{1}{2} \\ 1 \\ -2 \end{pmatrix}$ et son rayon vaut 3.

Fiche 12

12.1

1. Nous avons $A(X) = Q(X)B(X) + R(X)$ avec $Q(X) = X + 1$ et $R(X) = 0$.

2. Nous avons $A(X) = Q(X)B(X) + R(X)$ avec $Q(X) = X + 3$ et $R(X) = 4X + 6$.

3. Nous avons $A(X) = Q(X)B(X) + R(X)$ avec $Q(X) = 1$ et $R(X) = -4X + 4$.

12.2

1. Le polynôme P est de degré 4.

2. En utilisant les identités remarquables, nous avons $P(X) = (X^2 + 4X + 3)(X^2 - 4X - 1)$. Résoudre $P(X) = 0$ revient donc à résoudre $(X^2 + 4X + 3)(X^2 - 4X - 1) = 0$. Un produit de facteurs étant nul si l'un au moins des facteurs est nul, nous devons résoudre $X^2 + 4X + 3 = 0$ ou $X^2 - 4X - 1 = 0$. Les racines de $X^2 + 4X + 3$ sont -1 et -3 . Les racines de $X^2 - 4X - 1$ sont $2 + \sqrt{5}$ et $2 - \sqrt{5}$. Nous en déduisons que les racines de $P(X)$ sont $-1, -3, 2 + \sqrt{5}$ et $2 - \sqrt{5}$.

12.3 Si nous calculons $P(4)$, nous obtenons $P(4) = -64 + 96 - 36 + k = -4 + k$. 4 est donc racine de $P(X)$ lorsque $-4 + k = 0$ ou encore $k = 4$.

12.4

1. Le polynôme P est de degré 4.

2. Nous avons $P(-1) = (-1)^4 + (-1)^3 - 7(-1)^2 - 13(-1) - 6 = 0$. Nous en déduisons que -1 est racine du polynôme $P(X)$.

3. Nous avons $Q(X) = X^3 - 7X - 6$ (il suffit de poser la division euclidienne de $P(X)$ par $X+1$).

4. Le polynôme $Q(X)$ peut se factoriser sous la forme $(X+1)(X+2)(X-3)$. Les racines de $Q(X)$ sont les valeurs qui annulent ce polynôme. Il s'agit de $-1, -2$ et 3 .

12.5

1. Nous avons $P(-2) = 2(-2)^3 - 7(-2) + 2 = 0$. Nous en déduisons que -2 est racine de $P(X)$.

2. Nous avons $P(X) = (X+2)(2X^2 - 4X + 1)$. Les racines de $P(X)$ sont les valeurs de X qui annulent $P(X)$. Il s'agit ici de $-2, 1 + \frac{\sqrt{2}}{2}$ et $1 - \frac{\sqrt{2}}{2}$.

Fiche 13

13.1

$$1. F_1(X) = \frac{1}{X^2(X-1)} = -\frac{1}{X} - \frac{1}{X^2} + \frac{1}{X-1}$$

$$2. F_2(X) = \frac{X}{(X+1)^2(X-1)^3} = \frac{1}{16(X+1)} + \frac{1}{8(X+1)^2} - \frac{1}{16(X-1)} + \frac{1}{4(X-1)^3}$$

$$3. F_3(X) = \frac{X}{(X+3)(X^2+X+1)} = -\frac{3}{7(X+3)} + \frac{3X+1}{7(X^2+X+1)}$$

$$4. F_4(X) = \frac{1}{X^2-1} = \frac{1}{2(X-1)} - \frac{1}{2(X+1)}$$

$$5. F_5(X) = \frac{X+3}{X^4-5X^2+4} = -\frac{2}{3(X-1)} + \frac{1}{3(X+1)} + \frac{5}{12(X-2)} - \frac{1}{12(X+2)}$$

$$6. F_6(X) = \frac{X^4+X+1}{X^3-3X^2+2X} = X+3+\frac{1}{2X}-\frac{3}{X-1}+\frac{19}{2(X-2)}.$$

13.2

1. Comme $L_p(y'') = p^2 L_p(y) - py(0) - y'(0)$, $L_p(y') = p L_p(y) - y(0)$ et $L_p(\sin 2t) = \frac{2}{p^2+4}$, l'équation différentielle $y''(t) - y'(t) = \sin(2t)$ équivaut à $p^2 L_p(y) - py(0) - y'(0) - p L_p(y) + y(0) = \frac{2}{p^2+4}$. Comme $y(0) = y'(0) = 1$, nous obtenons finalement l'équation $p^2 L_p(y) - p - p L_p(y) = \frac{2}{p^2+4}$. En posant $Y = L_p(y)$, nous avons $p^2 Y - pY - p = \frac{2}{p^2+4}$.

2. À partir de l'équation obtenue à la question précédente, nous avons $Y(p^2 - p) = \frac{2}{p^2+4} + p = \frac{p^3+4p+2}{p^2+4}$ et donc $Y = \frac{p^3+4p+2}{p(p-1)(p^2+4)} = \frac{a}{p} + \frac{b}{p-1} + \frac{cp+d}{p^2+4}$ avec a, b, c et d des coefficients à déterminer. Nous obtenons $a = -\frac{1}{2}, b = \frac{7}{5}, c = \frac{1}{10}$ et $d = -\frac{2}{5}$. Nous obtenons donc la décomposition suivante $Y = -\frac{1}{2p} + \frac{7}{5(p-1)} + \frac{\frac{1}{10}p - \frac{2}{5}}{p^2+4}$.

3. À partir des formules d'inversion connues de la transformée de Laplace appliquées à la décomposition obtenue à la question précédente nous obtenons $y(t) = \frac{7}{5}e^t + \frac{1}{10} \cos 2t - \frac{1}{5} \sin 2t - \frac{1}{2}$.

Fiche 14

14.1 Les solutions sont de la forme $\{(x; \frac{5}{2}x + \frac{3}{2}z; z)\}$ avec $x, z \in \mathbb{R}$.

14.2 Les solutions sont de la forme $\{(\frac{4}{3}z; 0; z)\}$ avec $z \in \mathbb{R}$.

14.3 Pondérer cette équation revient à résoudre le système linéaire suivant :

$$(S) \begin{cases} 3x_1 = x_3 \\ 8x_1 = 2x_4 \\ 2x_2 = 2x_3 + x_4 \\ x_2 = 3x_3 \end{cases}$$

La solution de (S) est donnée par $\{(x_1; x_2; x_3; x_4)\} = \{(1; 5; 3; 4)\}$ (ou par tout vecteur de $\mathbb{R}^4$ proportionnel à $\{(1; 5; 3; 4)\}$).

14.4 Pondérer cette équation revient à résoudre le système linéaire suivant :

$$(S) \begin{cases} 2x_1 = x_3 \\ 3x_1 = x_4 \\ 2x_2 = 3x_3 + 2x_4 \\ x_2 = 3x_3 \end{cases}$$

La solution de (S) est donnée par $\{(x_1; x_2; x_3; x_4)\} = \{(1; 6; 2; 3)\}$ (ou par tout vecteur de $\mathbb{R}^4$ proportionnel à $\{(1; 6; 2; 3)\}$).

14.5

1. La solution de (S_1) est le triplet $\{(3; 5; 3)\}$.
2. La solution de (S_2) est le triplet $\{(1; -\frac{1}{2}; \frac{1}{2})\}$.
3. La solution de (S_3) est $\{(1; 2; -\frac{13}{3}; -\frac{2}{3})\}$.

Fiche 15

15.1 La solution du système (S) est le triplet $\{(1; 4; 4)\}$.

15.2 La solution du système (S) est le triplet $\{(3; 5; 3)\}$.

15.3 Pondérer cette équation revient à résoudre le système linéaire suivant :

$$(S) \begin{cases} 6x_1 = x_3 \\ 12x_1 = 2x_4 \\ 6x_1 + 2x_2 = 2x_3 + x_4 \end{cases}$$

La solution de (S) est donnée par $\{(x_1; x_2; x_3; x_4)\} = \{(1; 6; 6; 6)\}$ (ou par tout vecteur de $\mathbb{R}^4$ proportionnel à $\{(1; 6; 6; 6)\}$).

15.4 Pondérer cette équation revient à résoudre le système linéaire suivant :

$$(S) \begin{cases} 4x_1 = x_3 \\ 10x_1 = 2x_4 \\ 2x_2 = 2x_3 + x_4 \end{cases}$$

La solution de (S) est donnée par $\{(x_1; x_2; x_3; x_4)\} = \{(1; 4; 5; \frac{13}{2})\}$ (ou par tout vecteur de $\mathbb{R}^4$ proportionnel à $\{(1; 4; 5; \frac{13}{2})\}$).

15.5

1. La solution du système (S_1) est le triplet $\{(1; -\frac{1}{2}; \frac{1}{2})\}$.
2. La solution du système (S_2) est le triplet $\{(1; 2; 3)\}$.

Fiche 16

16.1 Sous forme algébrique, nous avons : $z_1 = 4 + 7i$, $z_2 = -16 + 30i$ et $z_3 = 34$.

16.2

1. Sous forme algébrique, nous avons $z_1 = \frac{1}{2} + i\frac{\sqrt{3}}{2}$ et $z_2 = \frac{\sqrt{2}}{2} - i\frac{\sqrt{2}}{2}$.

2. Sous forme algébrique, nous avons $z_1 z_2 = \frac{\sqrt{6} + \sqrt{2}}{4} + i\frac{\sqrt{6} - \sqrt{2}}{4}$. Sous forme exponentielle, nous avons, $z_1 z_2 = e^{i\frac{\pi}{12}}$. Enfin, sous forme trigonométrique, nous avons $z_1 z_2 = \cos \frac{\pi}{12} + i \sin \frac{\pi}{12}$.

3. Par identification entre les formes algébrique et trigonométrique de la question précédente, nous en déduisons que $\cos \frac{\pi}{12} = \frac{\sqrt{6} + \sqrt{2}}{4}$ et $\sin \frac{\pi}{12} = \frac{\sqrt{6} - \sqrt{2}}{4}$.

16.3

$$1. z = \frac{-11i - 10}{17}.$$

$$2. z = \frac{-20 + 5i}{17}.$$

16.4 $z^{12} = 1$. z est bien racine douzième de l'unité.

16.5

1. L'impédance complexe total du circuit est $Z_{tot} = \frac{iR(LC\omega^2 - 1)}{RC\omega + i(LC\omega^2 - 1)}$.

2. L'impédance est nulle lorsque $LC\omega^2 = 0$. De plus, $\omega = 2\pi f$ où f est la fréquence. Nous en déduisons que l'impédance est nulle pour $f = \frac{1}{2\pi\sqrt{LC}}$.

16.6 Nous avons, $(\sin \theta)^2 = \left(\frac{e^{i\theta} - e^{-i\theta}}{2i}\right)^2 = \frac{-1}{4}(e^{i\theta} - e^{-i\theta})^2 = \frac{-1}{4}(e^{i2\theta} + e^{-i2\theta} - 2) = \frac{1 - \cos(2\theta)}{2}$.

Fiche 17

17.1

1. Nous avons $z_A = 3\sqrt{2}\left(\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}i\right)$ et $z_B = 3\sqrt{2}\left(\frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}}i\right)$. Nous avons donc $|z_A| = |z_B| = 3\sqrt{2}$ et le triangle OAB est isocèle en O .

2. D'après la question précédente, nous avons $\arg(z_A) = \frac{\pi}{4}$ et $\arg(z_B) = -\frac{\pi}{4}$. Nous en déduisons que $(\vec{OA}; \vec{OB}) = \frac{\pi}{4} + \frac{\pi}{4} = \frac{\pi}{2}$.

3. Nous pouvons donc déduire des deux questions précédentes que le triangle OAB est rectangle et isocèle en O .

17.2 Nous avons $(\overrightarrow{CA}; \overrightarrow{CB}) = \frac{\pi}{3}$. Nous en déduisons que $\arg\left(\frac{z_A - z_C}{z_B - z_C}\right) = \frac{\pi}{3}$ et donc $\frac{a-c}{b-c} = e^{i\frac{\pi}{3}}$. Cette dernière équation permet d'écrire $a - be^{i\frac{\pi}{3}} + c(e^{i\frac{\pi}{3}} - 1) = 0$ ou encore $a + be^{i(\pi+\frac{\pi}{3})} + c\left(\frac{1}{2} + i\frac{\sqrt{3}}{2} - 1\right) = 0$. Nous en déduisons que $a + be^{i\frac{4\pi}{3}} + c\left(-\frac{1}{2} + i\frac{\sqrt{3}}{2}\right) = 0$ et donc $a + bj^2 + cj = 0$.

17.3

1. Il s'agit du cercle de centre le point $(0; 3)$ et de rayon 5.
2. Il s'agit du cercle de centre $(0; 0)$ et de rayon 2.
3. Nous avons $|z - z_A| = |z - z_B|$ ou encore $AM = BM$. L'ensemble des points M se trouve sur la médiatrice de $[AB]$.

17.4 Le fait que les points A, B et C soient alignés équivaut à ce que les vecteurs $\overrightarrow{AB}$ et $\overrightarrow{AC}$ soient colinéaires. Il existe donc un réel λ tel que $\overrightarrow{AB} = \lambda \overrightarrow{AC}$ donc $z_B - z_A = \lambda(z_C - z_A)$. Un autre raisonnement serait de dire que dans ce contexte, l'argument de $\frac{z_B - z_A}{z_C - z_A}$ est 0 modulo $k\pi$ avec $k \in \mathbb{Z}$. Cela nous permet de déduire que $\frac{z_B - z_A}{z_C - z_A}$ est un réel.

17.5 Le fait que les vecteurs $\overrightarrow{AB}$ et $\overrightarrow{AC}$ soient perpendiculaires équivaut à $(\overrightarrow{AB}; \overrightarrow{AC}) = \frac{\pi}{2}$ modulo $k\pi$ avec $k \in \mathbb{Z}$. Dans ce contexte, l'argument de $\frac{z_B - z_A}{z_C - z_A}$ est 0 modulo $k\pi$ avec $k \in \mathbb{Z}$. Cela nous permet de déduire que $\frac{z_B - z_A}{z_C - z_A}$ est un imaginaire pur.

Fiche 18

18.1

1. E_1 est un sous-espace vectoriel de $\mathbb{R}^3$.
2. E_2 n'est pas un sous-espace vectoriel de $\mathbb{R}^3$.

18.2 Tout d'abord la fonction nulle appartient à E_1 . D'autre part, si nous prenons $\lambda \in \mathbb{R}$, f et $g \in E_1$, alors, $\forall x \in \mathbb{R}$, $(\lambda f + g)(-x) = (\lambda f)(-x) + g(-x) = \lambda f(-x) + g(-x) = \lambda f(x) + g(x) = (\lambda f + g)(x)$. Donc E_1 est bien un sous-espace-vectoriel de E .

18.3 Tout d'abord, le polynôme nul est un polynôme de degré inférieur ou égal à trois. D'autre part, prenons P_1 et P_2 deux polynômes

de degré inférieur ou égal à trois et $\lambda \in \mathbb{R}$. Alors, λP_1 reste un polynôme de degré inférieur ou égal à trois et la somme de deux polynômes de degré inférieur ou égal à trois étant encore un polynôme de degré inférieur ou égal à trois, nous en déduisons que $\lambda P_1 + P_2$ est un polynôme de degré inférieur ou égal à trois.

18.4

1. Soient $\vec{u} = (x; y)$ et $\vec{v} = (x'; y') \in E$. Alors, $\vec{u} + \vec{v} = (x + x'; y + y')$. Comme x, x', y et $y' \geq 0$, en particulier $x + x' \geq 0$ et $y + y' \geq 0$. Nous en déduisons que $\vec{u} + \vec{v} \in E$.

2. Si nous choisissons un λ négatif, par exemple -1 et prenons $\vec{u} = (x; y) \in E$, alors $\lambda \vec{u} = (-x; -y) \notin E$ puisque $-x \leq 0$ et $-y \leq 0$. Cela suffit pour montrer que E n'est pas un $\mathbb{R}$ -espace vectoriel.

18.5

1. Soient $\lambda \in \mathbb{R}$ et $\vec{u} = (x; y) \in F$. Alors $\lambda \vec{u} = (\lambda x; \lambda y) \in F$ car $(\lambda x) \times (\lambda y) = \lambda^2 xy \geq 0$ puisque $xy \geq 0$.

2. Nous choisissons par exemple $\vec{u} = (1; 0)$ et $\vec{v} = (0; -1)$. Nous avons bien $\vec{u} \in F$ puisque $xy = 0 \geq 0$ et $\vec{v} \in F$ puisque $xy = 0 \geq 0$. Cependant, $\vec{u} + \vec{v} = (1; -1) \notin F$ puisque $xy = -1 < 0$. Cela suffit à montrer que F n'est pas un $\mathbb{R}$ -espace vectoriel.

Fiche 19

19.1

1. La famille E_1 forme une base de E .
2. La famille E_2 n'est pas libre donc ce n'est pas une base de E .
3. La famille E_3 n'est donc pas génératrice de E . Ce n'est donc pas une base de E .
4. La famille E_4 est libre et génératrice et forme ainsi une base de E .

19.2

1. La famille $\{P_1; P_2; P_3; P_4\}$ est libre. En effet, si nous supposons que

$$\forall X, \lambda_1 P_1(X) + \lambda_2 P_2(X) + \lambda_3 P_3(X) + \lambda_4 P_4(X) = 0,$$

alors, pour $X = 0$, nous avons $-6\lambda_4 = 0$ et donc $\lambda_4 = 0$. En posant $X = 1$, nous avons $6\lambda_2 = 0$ et

donc $\lambda_2 = 0$. Puis, en posant $X = 2$, nous avons $-2\lambda_3 = 0$ et donc $\lambda_4 = 0$. Nous en déduisons finalement que $\lambda_1 = 0$. Les quatre paramètres $\lambda_1, \lambda_2, \lambda_3$ et λ_4 étant nuls, nous en déduisons que la famille est libre.

2. Nous avons montré à la question précédente que la famille est libre. D'autre part, elle est composée de quatre éléments qui est la dimension de $\mathbb{R}_3[X]$. C'est donc une base de $\mathbb{R}_3[X]$.

19.3

1. Le vecteur $\begin{pmatrix} -3 \\ -1 \\ 1 \end{pmatrix}$ forme une base de E_1 .

2. Le vecteur $\begin{pmatrix} -1 \\ -1 \\ 1 \\ 0 \end{pmatrix}$ forme une base de E_2 .

3. Les vecteurs $\begin{pmatrix} 2 \\ 1 \\ 0 \\ 3 \end{pmatrix}$ et $\begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$ forment une base de E_3 .

19.4

1. Les vecteurs $\begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$ et $\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$ forment une base de F . F est de dimension trois.

2. Les vecteurs $\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -2 \\ 1 \\ 0 \end{pmatrix}$ et $\begin{pmatrix} 0 \\ -3 \\ 0 \\ 1 \end{pmatrix}$ forment une base de G . G est de dimension trois.

3. Les vecteurs $\begin{pmatrix} -4 \\ -2 \\ 1 \\ 0 \end{pmatrix}$ et $\begin{pmatrix} -6 \\ -3 \\ 0 \\ 1 \end{pmatrix}$ forment une base de $F \cap G$. $F \cap G$ est de dimension deux.

Fiche 20

20.1 Si nous utilisons les notations du Théorème 20.2, avec u et $v \in E$ et $\lambda \in \mathbb{K}$, $(f+g)(\lambda u+v) = f(\lambda u+v) + g(\lambda u+v) = \lambda f(u) + f(v) + \lambda g(u) + g(v) = \lambda(f+g)(u) + (f+g)(v)$. Nous en déduisons que la somme de deux applications linéaires est une application linéaire.

20.2 Si nous utilisons les notations du Théorème 20.2, avec u et $v \in E$ et $\lambda, \mu \in \mathbb{K}$, $(\lambda f)(\mu u+v) = \lambda f(\mu u+v) = \lambda f(\mu u) + \lambda f(v) = \mu(\lambda f)(u) + (\lambda f)(v)$. Nous en déduisons que le produit d'une application linéaire par un scalaire est une application linéaire.

20.3 Si nous utilisons les notations du Théorème 20.2, avec u et $v \in E$ et $\lambda \in \mathbb{K}$, $(h \circ g)(\lambda u+v) = h(g(\lambda u+v)) = h(\lambda g(u) + g(v)) = \lambda h(g(u)) + h(g(v)) = \lambda(h \circ g)(u) + (h \circ g)(v)$. Nous en déduisons que la composée de deux applications linéaires est une application linéaire.

20.4

1. Si nous choisissons par exemple $(x; y) = (0; 0)$ et $(x'; y') = (1; 0)$, alors $f(x; y) = 1$, $f(x'; y') = 2$ et $f((x; y) + (x'; y')) = 2$. Donc $f((x; y) + (x'; y')) \neq f(x; y) + f(x'; y')$ et f n'est pas une application linéaire.

2. Si nous choisissons par exemple $(x; y) = (1; 0)$ et $(x'; y') = (0; 1)$, alors $g(x; y) = 0$, $g(x'; y') = 0$ et $g((x; y) + (x'; y')) = 1$. Donc $g((x; y) + (x'; y')) \neq g(x; y) + g(x'; y')$ et g n'est pas une application linéaire.

20.5

1. Soient $x = x_1 + x_2$ et $x' = x'_1 + x'_2 \in E$ et $\lambda \in \mathbb{K}$. Alors $p(\lambda x + x') = p(\lambda x_1 + x'_1 + \lambda x_2 + x'_2) = \lambda x_1 + x'_1 = \lambda p(x) + p(x')$. Donc p est une application linéaire.

2. Soient $x = x_1 + x_2$ et $x' = x'_1 + x'_2 \in E$ et $\lambda \in \mathbb{K}$. Alors $s(\lambda x + x') = s(\lambda x_1 + x'_1 + \lambda x_2 + x'_2) = (\lambda x_1 + x'_1) - (\lambda x_2 + x'_2) = \lambda(x_1 - x_2) + (x'_1 - x'_2) = \lambda s(x) + s(x')$. Donc s est une application linéaire.

20.6 Soient $x, x' \in E$ et $\mu \in \mathbb{K}$. Alors $h_\alpha(\mu x + x') = \alpha(\mu x + x') = \mu(\alpha x) + \alpha x' = \mu h_\alpha(x) + h_\alpha(x')$. Nous en déduisons que h_α est une application linéaire. Cette application étant définie de E dans E , il s'agit d'un endomorphisme. Pour montrer que c'est un automorphisme, il reste à montrer qu'elle est bijective. Supposons que l'on ait x et $x' \in E$ tels que $h_\alpha(x) = h_\alpha(x')$. Alors $\alpha x = \alpha x'$ et donc $x = x'$. L'application est injective. D'autre part, si $y \in E$ tel que $y = h_\alpha(x)$, alors $y = \alpha x$ et il existe $x \in E$, $x = \frac{y}{\alpha}$ tel que x soit l'antécédent de y . Nous en déduisons que h_α est également

surjective. Elle est donc bijective et c'est donc un automorphisme.

20.7 Soient $u = (x_1; x_2)$ et $u' = (x'_1; x'_2) \in \mathbb{R}^2$ et $\lambda \in \mathbb{R}$. Alors $T(\lambda x + x') = T(\lambda x_1 + x'_1; \lambda x_2 + x'_2) = r(\lambda x_1 + x'_1; \lambda x_2 + x'_2) = \lambda r(x_1; x_2) + r(x'_1; x'_2) = \lambda T(u) + T(v)$. Donc T est une application linéaire.

Fiche 21

21.1

1. L'application n'est pas injective ni surjective et donc pas bijective.
2. L'application est injective mais n'est pas surjective. Elle n'est donc pas bijective.
3. L'application n'est pas injective ni surjective et donc pas bijective.
4. L'application n'est pas injective et donc pas bijective. Elle est cependant surjective.
5. L'application est injective. Elle n'est cependant pas surjective et donc pas bijective.

21.2

1. $f(u) = (x - y)e_1 + ze_3$.
2. $\text{Ker}(f) = \text{Vect}(1; 1; 0)$. Le théorème du rang permet d'en déduire que $\dim(\text{Im}(f)) = 2$ et $\text{Im}(f) = \text{Vect}((1; 0; 0); (0; 0; 1))$.
3. $\forall (x; y; z) \in \mathbb{R}^3, f \circ f(x; y; z) = f(f(x; y; z)) = f((x - y)e_1 + ze_3) = (x - y)f(e_1) + zf(e_3) = (x - y)e_1 + ze_3 = f(x; y; z)$. Nous en déduisons que $f \circ f = f$.

21.3

1. Nous avons $\text{Ker}(p) = G$. Le noyau n'étant pas réduit à l'élément neutre de E , l'application n'est pas injective.
2. Nous avons que le noyau de s_F est réduit à l'élément neutre de E . L'application est donc injective.

21.4

1. Nous avons $\text{Ker}(f) = \{k(X - 1)^2; k \in \mathbb{R}\}$. Le noyau n'étant pas réduit à l'élément neutre de E , l'application f n'est pas injective.
2. D'après le théorème du rang, $\dim(\text{Im}(f)) = 3 \neq \dim(\mathbb{R}_3[X])$, l'application f n'est donc pas surjective.

21.5 Nous avons montré dans la fiche précédente que c'est un endomorphisme. Le but ici est d'utiliser le noyau et l'image de cette application pour montrer qu'elle est bijective. Nous avons $\text{Ker}(f) = 0_E$ donc l'application est injective. D'après le théorème du rang, $\dim(\text{Im}(f)) = \dim(E)$ donc nous en déduisons que l'application est également surjective. L'application h_k étant à la fois injective et surjective, elle est bijective.

Fiche 22

22.1 La matrice B doit posséder 5 colonnes.

22.2 Les résultats sont $A = \begin{pmatrix} -1 & -9 \\ -2 & -9 \end{pmatrix}$ et $B = \begin{pmatrix} -1 & -6 \\ -2 & 5 \end{pmatrix}$.

22.3 Les résultats sont

$$\begin{aligned} -2A &= \begin{pmatrix} -4 & -6 & 0 & -2 \\ -2 & -10 & -2 & 0 \\ 0 & -4 & -6 & -4 \end{pmatrix}, \\ B - 2A &= \begin{pmatrix} -3 & -5 & 1 & -2 \\ -2 & -9 & 0 & 0 \\ 0 & -4 & -5 & -1 \end{pmatrix} \\ \text{et } A + 2B &= \begin{pmatrix} 4 & 5 & 2 & 1 \\ 1 & 7 & 5 & 0 \\ 0 & 2 & 5 & 8 \end{pmatrix}. \end{aligned}$$

Les produits AB et BA ne sont pas réalisables car dans chacun des cas le nombre de colonnes de la première matrice est différent du nombre de ligne de la seconde.

22.4 Le produit BA n'est pas réalisable. Par contre $AB = \begin{pmatrix} 5 & 11 & 17 & 23 \\ 11 & 25 & 39 & 53 \\ 17 & 39 & 61 & 83 \end{pmatrix}$.

22.5 Comme $A^2 = I$ où I est la matrice identité de taille 3, nous en déduisons que $A^{-1} = A$. Nous pouvons retrouver ce résultat en inversant directement A à partir de l'algorithme présenté dans la partie application de cette fiche.

22.6 $A^{-1} = \begin{pmatrix} 5 & -2 \\ -2 & 1 \end{pmatrix}, B^{-1} = \begin{pmatrix} 1 & -2 & 7 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}$
 et $C^{-1} = \begin{pmatrix} 1 & -4 & -3 \\ 1 & -5 & -3 \\ -1 & 6 & 4 \end{pmatrix}$.

Fiche 23

23.1

1. Les familles $\mathcal{B}$ et $\mathcal{B}'$ sont chacune composées de deux vecteurs non colinéaires. Nous en déduisons que ce sont des bases de $\mathbb{R}^2$.

2. Les coordonnées de $u = \begin{pmatrix} -1 \\ 5 \end{pmatrix}$ respectivement dans les bases $\mathcal{B}$ et $\mathcal{B}'$ sont $\begin{pmatrix} \frac{17}{11} \\ \frac{7}{11} \end{pmatrix}$ et $\begin{pmatrix} \frac{3}{27} \\ -\frac{21}{27} \end{pmatrix}$.

3. La matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$ est $\begin{pmatrix} \frac{1}{11} & -2 \\ -1 & \frac{14}{11} \end{pmatrix}$.

4. La matrice de passage de $\mathcal{B}'$ à $\mathcal{B}$ est $\begin{pmatrix} -\frac{34}{35} & \frac{22}{27} \\ -\frac{27}{14} & \frac{1}{27} \end{pmatrix}$.

23.2 Il s'agit du vecteur $\begin{pmatrix} -3 \\ 10 \\ 16 \end{pmatrix}$.

23.3 Il s'agit de la matrice $\begin{pmatrix} 1 & -1 & 0 \\ 1 & 2 & 0 \\ 3 & 0 & -4 \end{pmatrix}$.

23.4

1. Il s'agit de la matrice $\begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \end{pmatrix}$.

2. Il s'agit de la matrice $\begin{pmatrix} \frac{1}{3} & \frac{1}{3} & 0 \\ \frac{1}{12} & -\frac{1}{4} & 0 \end{pmatrix}$.

3. Il s'agit de la matrice $\begin{pmatrix} 0 & 0 \\ 2 & 2 \end{pmatrix}$.

4. Il s'agit de la matrice $\begin{pmatrix} \frac{2}{3} & \frac{2}{3} \\ -\frac{1}{3} & -\frac{1}{3} \end{pmatrix}$.

Fiche 24

24.1 $\begin{vmatrix} 1+a & a & a \\ b & 1+b & b \\ c & c & 1+c \end{vmatrix} = (1+a+b+c) \begin{vmatrix} 1 & 1 & 1 \\ b & 1+b & b \\ c & c & 1+c \end{vmatrix} = (1+a+b+c) \begin{vmatrix} 1 & 0 & 0 \\ b & 1 & 0 \\ c & 0 & 1 \end{vmatrix} = (1+a+b+c) \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1+a+b+c$.

24.2 La première colonne étant une combinaison linéaire des deuxième et troisième colonnes, le déterminant est nul.

24.3

1. Si nous développons selon les éléments de la première colonne, nous avons $\begin{vmatrix} 3 & 2 & 1 \\ 0 & 1 & 1 \\ 1 & 4 & 1 \end{vmatrix} = 3A_{11} + A_{31} = 3 \times -3 + 1 = -8$.

Si nous développons selon les éléments de la deuxième ligne, nous avons $\begin{vmatrix} 3 & 2 & 1 \\ 0 & 1 & 1 \\ 1 & 4 & 1 \end{vmatrix} = A_{22} + A_{23} = 2 - 10 = -8$. Nous retrouvons le même résultat qu'avec le développement par rapport à la première colonne.

2. Si nous développons selon les éléments de la première colonne, nous avons $\begin{vmatrix} 2 & -1 & -1 \\ -1 & 6 & 1 \\ 0 & 2 & -1 \end{vmatrix} = 2A_{11} - A_{21} = 2 \times -8 - 1 = -17$.

Si nous développons selon les éléments de la deuxième ligne, nous avons $\begin{vmatrix} 2 & -1 & -1 \\ -1 & 6 & 1 \\ 0 & 2 & -1 \end{vmatrix} = -A_{21} + 6A_{22} + A_{23} = -1 + 6 \times -2 - 4 = -17$. Nous retrouvons le même résultat qu'avec le développement par rapport à la première colonne.

24.4

1. $\begin{vmatrix} 1 & 2 & 4 \\ 0 & 1 & -2 \\ 5 & 6 & 2 \end{vmatrix} = -26$,

2. $\begin{vmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \\ 4 & 5 & 6 & 7 \end{vmatrix} = 0$ car la quatrième colonne est proportionnelle aux trois autres.

24.5 Il s'agit du même exercice que l'exemple donné en application de cette Fiche. Nous avons donc l'aire du parallélogramme qui est donnée par : $\mathcal{A} = \left| \det \begin{pmatrix} -1 & 3 \\ 4 & -5 \end{pmatrix} \right| = |5 - 12| = 7$.

24.6 Nous utilisons le dernier théorème de la partie cours. Le volume du parallélépipède est alors donné par : $\mathcal{V} = \left| \det \begin{pmatrix} 1 & 1 & 7 \\ 0 & 2 & 1 \\ -2 & 4 & 0 \end{pmatrix} \right| = |-4 + 26| = 22$.

Fiche 25

25.1

1. Le déterminant de la matrice A vaut -1 . Ce dernier étant non nul, la matrice A est inversible et $A^{-1} = \begin{pmatrix} 0 & 1 & -1 \\ 4 & -3 & 4 \\ 3 & -3 & 4 \end{pmatrix}$.

2. Le déterminant de la matrice B vaut -1 . Ce dernier étant non nul, la matrice B est inversible et $B^{-1} = \begin{pmatrix} 1 & -4 & -3 \\ 1 & -5 & -3 \\ -1 & 6 & 4 \end{pmatrix}$.

25.2

1. Le déterminant de la matrice A vaut $(\lambda + 2)(\lambda - 1)^2$. La matrice A est donc inversible lorsque ce déterminant est non nul, c'est-à-dire pour $\lambda \neq -2$ et $\lambda \neq 1$.

2. Dans le cas où $\lambda \neq -2$ et $\lambda \neq 1$, la matrice A est inversible et $A^{-1} = \frac{1}{(\lambda+2)(\lambda-1)^2} \begin{pmatrix} \lambda^2 - 1 & 1 - \lambda & 1 - \lambda \\ 1 - \lambda & \lambda^2 - 1 & 1 - \lambda \\ 1 - \lambda & 1 - \lambda & \lambda^2 - 1 \end{pmatrix}$.

25.3

1. Sous forme matricielle, le système s'écrit $AX = b$ avec $A = \begin{pmatrix} -1 & 1 & -2 \\ -1 & -1 & -1 \\ 2 & 1 & -1 \end{pmatrix}$, $b = \begin{pmatrix} -4 \\ -11 \\ 8 \end{pmatrix}$ et le vecteur des inconnues $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$.

Nous avons $\det(A) = -7 \neq 0$. Le système est donc un système de Cramer. La solution est donnée par le triplet $\{(3; 5; 3)\}$.

2. Sous forme matricielle, le système s'écrit $AX = b$ avec $A = \begin{pmatrix} 3 & 5 & -1 \\ 0 & 0 & 4 \\ 3 & 2 & 1 \end{pmatrix}$, $b = \begin{pmatrix} 2 \\ -6 \\ 1 \end{pmatrix}$ et le vecteur des inconnues $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$.

Nous avons $\det(A) = -4 \neq 0$. Le système est donc un système de Cramer. La solution est donnée par le triplet $\{(-\frac{13}{2}; 11; -\frac{3}{2})\}$.

25.4

1. Sous forme matricielle, le système s'écrit $AX = b$ avec $A = \begin{pmatrix} m & 1 & 1 & 1 \\ 1 & m & 1 & 1 \\ 1 & 1 & m & 1 \\ 1 & 1 & 1 & m \end{pmatrix}$, $b = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ et le vecteur des inconnues $X = \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix}$.

Nous avons $\det(A) = (m + 3)(m - 1)^2$. Il nous faut donc considérer divers cas de figure pour résoudre ce système.

2. Pour résoudre ce système, nous considérons plusieurs cas :

1. Si $m \neq -3$ et $m \neq 1$, le système est un système de Cramer et les solutions sont $x = \frac{m+2}{(m+3)(m-1)}$, $y = \frac{-1}{(m+3)(m-1)}$, $z = \frac{-1}{(m+3)(m-1)}$ et $t = \frac{-1}{(m+3)(m-1)}$.

2. Si $m = 1$ ou $m = -3$, il n'y a pas de solutions.

Fiche 26

26.1

1. Le polynôme caractéristique de A vaut $P_A(X) = (2 - X)(1 - X)$. Sachant que λ est valeur propre de A si elle annule son polynôme caractéristique, nous avons que 1 et 2 sont les valeurs propres de A .

2. Le vecteur propre associé à la valeur propre $\lambda = 1$ est $\begin{pmatrix} -2 \\ 3 \end{pmatrix}$. Celui associé à la valeur propre $\lambda = 2$ est $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

3. A possède deux valeurs propres distinctes, elle est donc diagonalisable c'est-à-dire qu'elle peut s'écrire sous la forme $A = PDP^{-1}$ avec $D = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ et $P = \begin{pmatrix} -2 & 1 \\ 3 & -1 \end{pmatrix}$.

26.2

1. Le polynôme caractéristique de B vaut $P_B(X) = (X + 1)^2$. Sachant que λ est valeur propre de B si elle annule son polynôme caractéristique, nous avons que -1 est valeur propre double de B .

2. Le vecteur propre associé à la valeur propre $\lambda = -1$ est $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

3. B possède une valeur propre double mais un seul vecteur propre associé à celle-ci. B n'est donc pas diagonalisable.

26.3

1. Nous avons $C \times \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \end{pmatrix} = 3 \times \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ et $C \times \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix} = -1 \times \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. Ce sont donc bien des vecteurs propres de C associés respectivement aux valeurs propres 3 et -1 .

2. C possède deux valeurs propres distinctes, elle est donc diagonalisable c'est-à-dire qu'elle peut s'écrire sous la forme $C = PDP^{-1}$ avec $D = \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix}$ et $P = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$.

26.4

1. Le polynôme caractéristique de E vaut $P_E(X) = X(1 - X)(X - 4)$. Sachant que λ est valeur propre de E si elle annule son polynôme caractéristique, nous avons que 0, 1 et 4 sont les valeurs propres de E .

2. Le vecteur propre associé à la valeur propre $\lambda = 0$ est $\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$. Celui associé à la valeur propre

$\lambda = 1$ est $\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$. Enfin, celui associé à la valeur propre $\lambda = 4$ est $\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$.

3. E possède trois valeurs propres distinctes, elle est donc diagonalisable c'est-à-dire qu'elle

peut s'écrire sous la forme $E = PDP^{-1}$ avec $D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix}$ et $P = \begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & 2 \\ -1 & 1 & 1 \end{pmatrix}$.

26.5

1. Le polynôme caractéristique de F vaut $P_F(X) = (1 - X)^3$. Sachant que λ est valeur propre de F si elle annule son polynôme caractéristique, nous avons que 1 est valeur propre triple de F .

2. Les vecteurs propres associés à la valeur propre $\lambda = 1$ sont $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ et $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.

3. F possède une valeur propre triple mais seulement deux vecteurs propres associés. Elle n'est donc pas diagonalisable.

Fiche 27

27.1 Nous utilisons le raisonnement par récurrence. Tout d'abord, pour $p = 0$, nous avons $A^p = A^0 = I_n$ et $PD^pP^{-1} = PP^{-1} = I_n$. L'égalité est donc bien vérifiée. Supposons maintenant que $A^p = PD^pP^{-1}$ à un certain ordre $p \in \mathbb{N}$ fixé. Montrons que nous avons l'égalité à l'ordre $p + 1$. Nous avons, pour $p \in \mathbb{N}$, $A^{p+1} = AA^p = APD^pP^{-1} = PDP^{-1}PD^pP^{-1} = PDD^pP^{-1} = PD^{p+1}P^{-1}$. L'égalité est donc bien vérifiée à l'ordre $p + 1$. Elle est donc vraie pour tout $p \in \mathbb{N}$.

27.2 Nous avons, pour tout $n \in \mathbb{N}^*$,

$$A^n = \begin{pmatrix} \frac{3^n + (-1)^n}{2} & \frac{3^n - (-1)^n}{2} \\ \frac{3^n - (-1)^n}{2} & \frac{3^n + (-1)^n}{2} \end{pmatrix}.$$

27.3 Nous avons, pour tout $n \in \mathbb{N}^*$,

$$B^n = \frac{1}{12} \begin{pmatrix} 4 + 2 \times 4^n & 3 \times 4^n & -4 + 7 \times 4^n \\ 16 + 2 \times 4^n & 3 \times 4^n & -16 + 7 \times 4^n \\ -8 + 2 \times 4^n & 3 \times 4^n & 8 + 7 \times 4^n \end{pmatrix}.$$

27.4 Nous avons, pour $t \in \mathbb{R}$, $x_1(t) = y_1(t) - y_2(t) + y_3(t)$, $x_2(t) = y_1(t) + y_2(t)$ et $x_3(t) = y_1(t) + y_3(t)$ avec $y_1(t) = k_1 e^{2t} + \frac{1}{2}(t^2 - 1)$, $y_2(t) = k_2 e^t - t^2 - 2t - 1$ et $y_3(t) = k_3 e^t - 2t^2 - 3t - 2$ où k_1 , k_2 et k_3 sont des constantes réelles.

27.5 Nous avons, pour $t \in \mathbb{R}$, $x_1(t) = (2t + 1)e^t - e^{2t}$ et $x_2(t) = (2t + 2)e^t - 2e^{2t}$.

27.6 Pour $n \in \mathbb{N}^*$, $u_n = \frac{1}{3}(2^n + 2(-1)^n)$.

Fiche 28

28.1 Tout d'abord, f est bilinéaire. En effet, soient $\lambda, \mu \in \mathbb{R}$, $x = (x_1; x_2)$, $y = (y_1; y_2)$, $z = (z_1; z_2)$ et $t = (t_1; t_2) \in \mathbb{R}^2$. Alors $f(\lambda x + z; \mu y + t) = (\lambda x_1 + z_1)(\mu y_2 + t_2) + (\lambda x_2 + z_2)(\mu y_1 + t_1) = \lambda\mu(x_1y_2 + x_2y_1) + \lambda(x_1t_2 + x_2t_1) + \mu(z_1y_2 + z_2y_1) + (z_1t_2 + z_2t_1) = \lambda\mu f(x; y) + \lambda f(x; t) + \mu f(z; y) + f(z; t)$.

Ensuite, elle est symétrique. En effet, $f(y; x) = y_1x_2 + y_2x_1 = f(x; y)$. Nous avons donc une forme bilinéaire symétrique.

28.2 f n'est pas une forme bilinéaire. En effet, si nous choisissons $x = (1; 0; 1)$, $y = (0; 1; 1)$ et $z = (0; 0; 1) \in \mathbb{R}^3$, nous avons en particulier que $f(x; y + z) = 4 \neq 2 = f(x; y) + f(x; z)$. Cela suffit à montrer que f n'est pas bilinéaire.

28.3

1. Montrons d'abord que f est bilinéaire. Soient $\lambda, \mu \in \mathbb{R}$, $F, G, H, R \in E$. Alors, $f(\lambda F + H; \mu G + R) = \int_0^1 (\lambda F + H)(t)(\mu G + R)(t)dt = \lambda\mu \int_0^1 F(t)G(t)dt + \lambda \int_0^1 F(t)R(t)dt + \mu \int_0^1 H(t)G(t)dt + \int_0^1 H(t)R(t)dt = \lambda\mu f(F; G) + \lambda f(F; R) + \mu f(H; G) + f(H; R)$. Nous en déduisons que f est bien bilinéaire. D'autre part, elle est symétrique puisque $f(F; G) = \int_0^1 F(t)G(t)dt = \int_0^1 G(t)F(t)dt = f(G; F)$. Enfin, elle est définie positive puisque $f(F, F) = \int_0^1 F^2(t)dt > 0$ si F est non nulle.

2. Nous avons $q(F) = f(F; F) = \int_0^1 F^2(t)dt$.

28.4 Nous devons pour cela montrer que $\langle P|Q \rangle$ est une forme bilinéaire symétrique définie positive sur E . Tout d'abord, soient $\lambda, \mu \in \mathbb{R}$, $P, Q, R, T \in \mathbb{R}_2[X]$. Nous avons $\langle \lambda P + R; \mu Q + T \rangle = (\lambda P + R)(-1)(\mu Q + T)(-1) + (\lambda P + R)(0)(\mu Q + T)(0) + (\lambda P + R)(1)(\mu Q + T)(1) = \lambda\mu(P(-1)Q(-1) + P(0)Q(0) + P(1)Q(1)) + \lambda(P(-1)T(-1) + P(0)T(0) + P(1)T(1)) + \mu(R(-1)Q(-1) + R(0)Q(0) + R(1)Q(1)) + (R(-1)T(-1) + R(0)T(0) + R(1)T(1)) = \lambda\mu \langle P|Q \rangle + \lambda \langle P|T \rangle + \mu \langle R, Q \rangle + \langle R, T \rangle$. Nous sommes bien donc en présence d'une forme bilinéaire.

Cette forme bilinéaire est symétrique. En effet, $\langle P, Q \rangle = P(-1)Q(-1) + P(0)Q(0) + P(1)Q(1) = Q(-1)P(-1) + Q(0)P(0) + Q(1)P(1) = \langle Q, P \rangle$.

Montrons enfin qu'elle est définie positive. Prenons $P(X) = a_0 + a_1X + a_2X^2$ un polynôme de $\mathbb{R}_2[X]$ non nul c'est à dire que les constantes a_0, a_1 et a_2 ne sont pas nulles simultanément. Alors, après calcul, $\langle P, P \rangle = 3a_0^2 + 2a_1^2 + 2a_2^2 + 4a_0a_2$ qui ne sera jamais nul compte tenu de la remarque précédente sur les coefficients a_0, a_1 et a_2 . Nous en déduisons que $\langle P, Q \rangle$ est une forme bilinéaire symétrique définie positive. C'est donc une norme.

Fiche 29

29.1

1. Si C_1, C_2 et C_3 sont respectivement la première, deuxième et troisième colonne de la matrice M , nous avons $\langle C_1|C_2 \rangle = \frac{1}{81}(-8 - 8 + 16) = 0$, $\langle C_1|C_3 \rangle = \frac{1}{81}(-32 + 4 + 28) = 0$ et $\langle C_2|C_3 \rangle = \frac{1}{81}(4 - 32 + 28) = 0$. Nous en déduisons que la matrice M est orthogonale.

2. Comme la matrice M est orthogonale, c'est une isométrie vectorielle. De plus, elle est symétrique, c'est donc une symétrie vectorielle. D'autre part, si I est la matrice identité de taille 3, nous avons $(M - I) \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ qui équivaut à $x + y + 4z = 0$. f est donc une réflexion par rapport au plan (P) d'équation $x + y + 4z = 0$.

29.2 La nouvelle base orthonormale est $\{v'_1; v'_2; v'_3\}$ avec $v'_1 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, $v'_2 = \frac{1}{\sqrt{6}} \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}$

$$\text{et } v'_3 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}.$$

29.3 Les matrices de rotation respectivement autour de x, y et z sont

$$R_x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix}, R_y = \begin{pmatrix} \cos \theta & 0 & \sin \theta \\ 0 & 1 & 0 \\ -\sin \theta & 0 & \cos \theta \end{pmatrix}$$

$$\text{et } R_z = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

29.4 Il s'agit de la matrice $\begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}$ avec $\theta = \frac{\pi}{2}$, soit $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

Fiche 30

30.1 Soit le point $M \begin{pmatrix} x \\ y \end{pmatrix}$ appartenant à la parabole. Ses coordonnées vérifient $\left(x - \frac{1}{2}\right)^2 + (y - 2)^2 = (x - 3)^2$. L'équation cartésienne de la parabole est $4y^2 - 16y + 20x - 19 = 0$.

30.2 C'est une parabole de sommet S qui a pour coordonnées $S \begin{pmatrix} -\frac{4}{5} \\ -\frac{16}{5} \end{pmatrix}$ et d'axes $\frac{1}{5} \begin{pmatrix} 4 \\ -3 \end{pmatrix}$ et $\frac{1}{5} \begin{pmatrix} 3 \\ 4 \end{pmatrix}$.

30.3
1. L'équation de la courbe peut s'écrire sous la forme $\frac{(x + \frac{4}{3})^2}{\frac{4}{9}} - \frac{y^2}{\frac{4}{3}} = 1$. Il s'agit donc d'une hyperbole d'équation $\frac{X^2}{a^2} - \frac{Y^2}{b^2} = 1$ avec $a^2 = \frac{4}{9}$, $b^2 = \frac{4}{3}$.

2. Le foyer F a pour coordonnées $F \begin{pmatrix} \frac{4}{3} \\ 0 \end{pmatrix}$. La directrice $\mathcal{D}$ a pour équation $x = \frac{1}{3}$ et l'excentricité e vaut 2.

30.4 Il s'agit d'une ellipse d'équation $\frac{(x-1)^2}{10} + \frac{(y+3)^2}{\frac{20}{3}} = 1$. Si nous posons $a = \sqrt{10}$, $b = \sqrt{\frac{20}{3}}$ et $c = \sqrt{a^2 - b^2}$, nous avons l'excentricité $e = \frac{c}{a}$, la directrice $\mathcal{D}$ qui a pour équation $x = \frac{a^2}{c}$ et le foyer F qui a pour coordonnées $F \begin{pmatrix} c \\ 0 \end{pmatrix}$.

30.5
1. Pour $\lambda = 1$ ou -1 , C_λ est une parabole. Pour $|\lambda| > 1$, C_λ est une hyperbole. Enfin, pour $|\lambda| < 1$, C_λ est une ellipse.

2. Pour $\lambda = 0$, nous avons $x^2 + y^2 = 1$. C_0 représente donc le cercle unité. Pour $\lambda = 1$, nous avons $(x+y)^2 = 1$. C_1 représente donc les droites d'équation $y = 1 - x$ et $y = -1 - x$. Enfin, pour $\lambda = -1$, nous avons $(x-y)^2 = 1$. C_{-1} représente donc les droites d'équation $y = x - 1$ et $y = x + 1$.

Fiche 31

31.1 Il faut que $x \geq 0$ et $E(x) \geq 0$ donc x est un réel positif ou nul. Posons $n = \sqrt{E(x)} \Rightarrow n^2 = E(x)$ et $x \in [n^2; n^2 + 1[$, donc si x est solution, il existe un entier naturel n tel que $x \in [n^2; n^2 + 1[$. Réciproquement, s'il existe un entier naturel n tel que $x \in [n^2; n^2 + 1[$, alors $E(x) = n^2$ et $x \in [n^2; \underbrace{n^2 + 2n + 1}_{(n+1)^2}]$ car $n^2 + 1 \leq n^2 + 2n + 1$.

Donc $\sqrt{x} \in [n; n+1[$ et $E(\sqrt{x}) = n = \sqrt{E(x)}$, x est solution. L'ensemble des solutions est l'ensemble : $\bigcup_{n \in \mathbb{N}} [n^2; n^2 + 1[$.

31.2 Posons $k = E(x)$. Il existe alors $\alpha \in [0; 1[$ tel que : $x = k + \alpha$. Ainsi $nx = nk + n\alpha$. De même posons $p = E(n\alpha)$. Il existe alors $\beta \in [0; 1[$ tel que : $n\alpha = p + \beta$. Ainsi $nx = nk + p + \beta$ avec $nk + p \in \mathbb{N}$ et $\beta \in [0; 1[$. Ainsi $E(nx) = nk + p$ donc $\frac{1}{n}E(nx) = k + \frac{p}{n}$ où $0 \leq \frac{p}{n} \leq \alpha < 1$ donc $k = E\left(\frac{1}{n}E(nx)\right) = E(x)$.

31.3 Nous avons, pour tout $x \in \mathbb{R}$ $E(x) \leq x < E(x) + 1$. Distinguons deux cas.

Premier cas : si $E(x)$ est pair. Posons $E(x) = 2p$, avec $p \in \mathbb{Z}$. Nous pouvons alors écrire : $2p \leq x < 2p + 1$, ce qui entraîne $p + \frac{1}{2} \leq \frac{x}{2} + \frac{1}{2} < p + 1$. Par conséquent $E\left(\frac{x}{2}\right) = p$ et $E\left(\frac{x+1}{2}\right) = p$ et donc nous avons $E\left(\frac{x}{2}\right) + E\left(\frac{x+1}{2}\right) = p + p = 2p = E(x)$.

Second cas : si $E(x)$ est impair. Posons $E(x) = 2p + 1$, avec $p \in \mathbb{Z}$. Nous pouvons alors écrire : $2p + 1 \leq x < 2p + 2$, ce qui entraîne $p + 1 \leq \frac{x}{2} + \frac{1}{2} < p + \frac{3}{2}$. Par conséquent

$$E\left(\frac{x}{2}\right) + E\left(\frac{x+1}{2}\right) = p + p + 1 = 2p + 1 = E(x)$$

et donc nous avons $E\left(\frac{x}{2}\right) + E\left(\frac{x+1}{2}\right) = E(x)$.

31.4 $\sup(F_1) = 0$, $\sup(F_2) = \frac{1}{2}$, $\inf(F_1) = -\ln 2$ et $\inf(F_2) = -\frac{1}{2}$. $\inf(F_3) = 2$. F_3 n'admet pas de borne supérieure.

Fiche 32

32.1 Soit la propriété : $\mathcal{P}(n) : \forall n \geq 2$, $u_n = \frac{1}{2}$. Nous allons démontrer cette propriété par récurrence sur l'entier n .

Initialisation : $u_2 = u_1^2 + \frac{1}{4} = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$. Donc $\mathcal{P}_2$ est vraie.

Hérédité : supposons $\mathcal{P}_n$ vraie pour un certain entier naturel $n \geq 2$. Montrons que $\mathcal{P}(n+1)$ est encore vraie. Nous avons $u_{n+1} = u_n^2 + \frac{1}{4} = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$, d'après l'hypothèse de récurrence, à savoir $u_n = \frac{1}{2}$. Donc nous en déduisons que $\mathcal{P}(n+1)$ est vraie.

Conclusion : nous concluons par récurrence sur n que pour tout entier $n \geq 2$, la suite (u_n) est constante. Donc la suite (u_n) est stationnaire.

32.2 La suite (u_n) est strictement croissante. Pour le montrer, vous pouvez calculer, par exemple la différence $u_{n+1} - u_n$. Après calculs, vous obtenez $u_{n+1} - u_n = \frac{2e^n(e-1)}{(e^{n+1}+1)(e^n+1)}$.

32.3 La suite (u_n) est strictement croissante. Pour le montrer, vous pouvez calculer, par exemple la différence $u_{n+1} - u_n$. Après calculs, vous obtenez $u_{n+1} - u_n = -2\left(\frac{2}{3}\right)^n\left(-\frac{1}{3}\right)$, différence qui est bien positive.

Fiche 33

33.1 La suite (u_n) définie par $u_n = n^3 + 7$ est divergente car $\lim_{n \rightarrow +\infty} u_n = +\infty$.

33.2

- La suite de terme général $t_n = \frac{1}{n^2}$ est négligeable devant la suite de terme général $u_n = \frac{1}{n}$.

- La suite de terme général $v_n = \frac{1}{2n}$ est dominée par la suite de terme général $w_n = \frac{1}{n}$.

- La suite de terme général $x_n = \frac{1}{n}$ est équivalente à la suite de terme général $y_n = \frac{1}{n-1}$.

33.3 Ces deux suites sont adjacentes. En effet, (u_n) est décroissante, (v_n) est croissante et $\lim_{n \rightarrow +\infty} (v_n - u_n) = 0$, car $\lim_{n \rightarrow +\infty} (v_n - u_n) = \lim_{n \rightarrow +\infty} -\frac{2}{n} = 0$. De plus, (u_n) et (v_n) convergent vers la même limite l qui est égale à 3.

33.4

- Comme il a été montré que la suite $(u_n)_{n \in \mathbb{N}^*}$ définie par $u_n = 7 + \frac{3}{n}$ est strictement décroissante et minorée par 7, nous en déduisons grâce au théorème du cours, que cette suite converge.

- Nous avons montré que $(u_n)_{n \in \mathbb{N}^*}$ définie par le terme général $u_n = \frac{1 \cdot 1! + 2 \cdot 2! + \dots + n \cdot n!}{(n+1)!} = 1 - \frac{1}{(n+1)!}$ est strictement croissante. Montrons que $(u_n)_{n \in \mathbb{N}^*}$ est majorée. La suite $(u_n)_{n \in \mathbb{N}^*}$ est majorée par 1 car nous retirons à 1 une quantité strictement positive. Par conséquent $(u_n)_{n \in \mathbb{N}^*}$ converge car $(u_n)_{n \in \mathbb{N}^*}$ est strictement croissante et majorée par 1.

33.5 La suite (u_n) est convergente. En effet, $\lim_{n \rightarrow +\infty} u_n = 2$. La suite (v_n) est divergente. En effet, $\lim_{n \rightarrow +\infty} v_n = +\infty$.

33.6 La suite (u_n) est convergente et $\lim_{n \rightarrow +\infty} u_n = 0$. La suite (v_n) est convergente et $\lim_{n \rightarrow +\infty} v_n = 1$. La suite (w_n) est divergente. Pour tout $n \in \mathbb{N}$, $\cos(n\pi) = (-2)^n$.

33.7

- La suite (u_n) est strictement croissante. En effet, en calculant la différence entre deux termes consécutifs, nous obtenons $u_{n+1} - u_n = \frac{1}{(n+1)^{n+1}} > 0$.
- D'autre part, notons, que pour tout $n \in \mathbb{N}$, $u_n \leq \lim_{n \rightarrow +\infty} \sum_{k=0}^n \frac{1}{2^k} = 2$. Donc tous les termes de la suite (u_n) sont majorés par 2.

Ainsi la suite (u_n) est une suite strictement croissante et majorée. Donc la suite (u_n) est convergente.

Fiche 34

34.1 La suite (u_n) est une suite géométrique de raison $q = 5 > 1$ et de premier terme 2. Nous avons alors : $\lim_{n \rightarrow +\infty} u_n = \lim_{n \rightarrow +\infty} 2 \times 5^n = +\infty$. La suite (u_n) diverge.

34.2

1. $v_{n+1} = 5u_{n+1} + 3 = 5 \frac{u_n - 3}{6} + 3 = \frac{5u_n + 3}{6} = \frac{1}{6}v_n$. Donc la suite (v_n) est une suite géométrique de raison $\frac{1}{6}$. Déduisons maintenant une expression de u_n en fonction de n . Pour cela, déterminons une expression de v_n en fonction de n . Pour tout $n \in \mathbb{N}$, nous avons $v_n = \frac{1}{6^n}v_0 = \frac{1}{6^n} \left(5 \frac{3}{5} + 3 \right) = \frac{1}{6^{n-1}}$. Finalement, pour tout $n \in \mathbb{N}$, $u_n = \frac{1}{5} \left(\frac{1}{6^{n-1}} - 3 \right)$.

2. $u_{n+1} - u_n = \frac{1}{5} \times \frac{1}{6^{n-1}} \times \frac{-5}{6} < 0$. Donc (u_n) est strictement décroissante. De plus, lorsque $n \rightarrow +\infty$, $\frac{1}{6^{n-1}} \rightarrow 0$ et ainsi $u_n \rightarrow -\frac{3}{5}$. Donc (u_n) est convergente.

$$3. S_n = v_0 + \dots + v_{n-1} = 6 \sum_{i=0}^{n-1} \frac{1}{6^i} = 6 \frac{1 - \left(\frac{1}{6}\right)^n}{1 - \frac{1}{6}} =$$

$$\frac{36}{5} \left(1 - \left(\frac{1}{6}\right)^n \right).$$

$$S'_n = \frac{v_0}{5} - \frac{3}{5} + \dots + \frac{v_{n-1}}{5} - \frac{3}{5} = \frac{1}{5} \left(\sum_{i=0}^{n-1} v_i \right) -$$

$$n \frac{3}{5} = \frac{36}{25} \left(1 - \left(\frac{1}{6}\right)^n \right) - n \frac{3}{5}.$$

Lorsque $n \rightarrow +\infty$, $\left(\frac{1}{6}\right)^n \rightarrow 0$ et ainsi $S_n \rightarrow \frac{36}{5}$ et $S'_n \rightarrow -\infty$.

34.3

1. La suite (v_n) est une suite géométrique de raison $\frac{1}{3}$. En effet, après calcul, nous montrons que $v_{n+1} = 4u_{n+1} - 6(n+1) + 15 = \frac{1}{3}(4u_n - 6n + 15) = \frac{1}{3}v_n$. Ainsi, nous en déduisons que pour tout $n \in \mathbb{N}$, $v_n = \frac{19}{3^n}$.

2. D'après la définition du terme général v_n , nous pouvons déduire que le terme général u_n est égal à $\frac{v_n}{4} + \frac{6n-15}{4} = \frac{19}{4 \times 3^n} + \frac{6n-15}{4}$.

3. En regardant l'expression du terme général u_n , nous pouvons établir que u_n peut se décomposer en deux termes t_n et w_n où $t_n = \frac{19}{4} \times \frac{1}{3^n}$ et $w_n = \frac{6n-15}{4}$. Ainsi la suite (t_n) est une suite géométrique de premier terme $t_0 = \frac{19}{4}$ et de raison $\frac{1}{3}$ et la suite (w_n) est une suite arithmétique de premier terme $w_0 = -\frac{15}{4}$ et de raison $\frac{3}{2}$.

$$4. T_n = \frac{19}{8} \left(3 - \left(\frac{1}{3}\right)^n \right), W_n = (n+1) \left(-\frac{15}{4} + \frac{3n}{4} \right) \text{ et } U_n = \frac{19}{8} \left(3 - \left(\frac{1}{3}\right)^n \right) + (n+1) \times \left(-\frac{15}{4} + \frac{3n}{4} \right).$$

34.4

$$1. u_1 = \frac{u_0 + v_0}{2} = \frac{7}{2}, v_1 = \frac{u_1 + v_0}{2} = \frac{15}{4}, u_2 = \frac{u_1 + v_1}{2} = \frac{29}{8} \text{ et } v_2 = \frac{u_2 + v_1}{2} = \frac{59}{16}.$$

2. $w_{n+1} = v_{n+1} - u_{n+1} = \frac{u_n + v_n}{2} - \frac{u_{n+1} + v_n}{2} = \frac{u_n - v_n}{4} = \frac{w_n}{4}$. Puis $w_0 = v_0 - u_0 = 4 - 3 = 1$. Ainsi $w_n = 4^{-n}w_0 = 4^{-n}$, pour tout $n \in \mathbb{N}$. Par conséquent, w_n tend vers 0 lorsque n tend vers $+\infty$.

3. $u_{n+1} - u_n = \frac{v_n - u_n}{2} = \frac{w_n}{2} = \frac{4^{-n}}{2} \geq 0$. Donc (u_n) est croissante.

$v_{n+1} - v_n = \frac{u_n - v_n}{4} = -\frac{w_n}{4} = -4^{-n-1} \geq 0$. Donc (v_n) est décroissante.
 $|u_n - v_n| = |w_n| = 4^{-n} \rightarrow 0$ quand $n \rightarrow +\infty$.
 Les deux suites (u_n) et (v_n) sont adjacentes, elles convergent vers la même limite L .

$$\begin{aligned} 4. j_{n+1} &= \frac{u_{n+1} + 2v_{n+1}}{3} = \frac{u_n + v_n + 2u_{n+1} + 2v_n}{6} \\ &= \frac{2u_n + 4v_n}{6} = j_n = \dots = j_0 = \frac{11}{3}. \\ j_n &\rightarrow \frac{L + 2L}{3} = L \text{ quand } n \rightarrow +\infty. \text{ Donc, par unicité de la limite, } L = \frac{11}{3}. \end{aligned}$$

Fiche 35

35.1 La suite (u_n) est une suite stationnaire. En effet, nous montrons pour tout $n \geq 2$, $u_n = \frac{1}{2}$.

35.2

1. En calculant la dérivée de f , qui existe puisque c'est une fonction rationnelle dont le dénominateur ne s'annule pas sur I , nous trouvons $f'(x) = \frac{6}{(x+4)^2}$. Donc la fonction f est strictement croissante sur I . Donc, pour tout $x \in I$, $f(0) = \frac{1}{2} \leq f(x) \leq f(1) = 1$. Donc, pour tout $x \in I$, $f(x)$ appartient à I .

2. En raisonnant par récurrence sur n , nous montrons que pour tout $n \in \mathbb{N}$, $u_n \in I$.

3. En raisonnant par récurrence sur n , nous montrons que la suite (u_n) est strictement croissante, en utilisant le fait que la fonction f est strictement croissante sur I .

4. La suite (u_n) est majorée et croissante, donc (u_n) converge vers $\ell \in I$.

5. Cette limite ℓ est telle qu'elle vérifie l'équation $\ell = f(\ell)$, car f est continue sur I . En résolvant l'équation $f(x) = x$, nous obtenons deux solutions réelles dont la seule acceptable est égale à 1. Donc la limite ℓ est égale à 1.

Fiche 36

36.1 Pour tout $x \in \mathbb{R}$, $E(2-x) = E(-x) + 2$. Donc $f(x) = 2 + E(x) + E(-x)$. Ainsi, pour tout $x \in \mathbb{R}$, $f(-x) = f(x)$ et $f(1+x) = f(x)$. Donc la fonction f est paire et a 1 pour période.

Par conséquent, il suffit d'étudier la continuité de f sur $\left[0; \frac{1}{2}\right]$. Nous avons pour tout $x \in$

$\left[0; \frac{1}{2}\right]$, $E(x) = 0, E(-x) = -1$, donc $f(x) = 1$ or $f(0) = 2$. Donc pour tout $x \in \mathbb{Z}$, $f(x) = 2$ et pour tout $x \in \mathbb{R} \setminus \mathbb{Z}$, $f(x) = 1$. La fonction f est donc continue en tout point de $\mathbb{R} \setminus \mathbb{Z}$ et discontinue en tout point de $\mathbb{Z}$.

36.2 Pour tout $x \in \mathbb{R}$, $E(x) < x$ donc le domaine de définition est $\mathbb{R}$. La fonction partie entière est continue sur tout intervalle $[n; n+1[$, $n \in \mathbb{Z}$ car elle est constante sur cet intervalle. La fonction $x \mapsto \sqrt{x}$ est continue sur $\mathbb{R}^+$ donc f est continue sur $[n; n+1[$. Nous étudions la continuité en $n \in \mathbb{Z}$: $f(n) = n$. Soit $x \in]n; n+1[$, $f(x) = E(x) - \sqrt{x - E(x)} = n + \sqrt{x - n}$, et par conséquent $\lim_{x \rightarrow n^+} f(x) = n = f(n)$. Donc f est continue à droite en n . Soit $x \in]n-1; n[$, $f(x) = n-1 + \sqrt{x - n + 1}$, $\lim_{x \rightarrow n^-} \sqrt{x - n + 1} = 1$ et par conséquent $\lim_{x \rightarrow n^-} f(x) = n = f(n)$. Donc f est continue à gauche en n . Ainsi f est continue en n et donc f est continue sur $\mathbb{R}$.

36.3 Pour tout $x \in \mathbb{R}$, $f(x+1) = f(x) + 1$. Il suffit donc d'étudier f sur l'intervalle $[0; 1]$. Pour tout $x \in [0; 1[$, $f(x) = x^2$ et $f(1) = 1$. Nous avons $\lim_{x \rightarrow 1^-} f(x) = 1$. La fonction f est donc continue sur l'intervalle $[0; 1]$ et continue à gauche au point 1. Donc f est continue sur $\mathbb{R}$.

36.4 Pour $-1 < a < 1$, $\lim_{x \rightarrow a} f(x) = 1 - \sqrt{1 - a^2} = f(a)$. Or, $\lim_{x \rightarrow -1^+} f(x) = 1 = f(-1)$ et $\lim_{x \rightarrow 1^-} f(x) = 1 = f(1)$: f est continue à droite en -1 et à gauche en 1 . La fonction f est continue sur $[-1; +1]$.

36.5 Soit $x \in]-1; 1[$, nous avons alors $\lim_{n \rightarrow +\infty} x^{2n} = 0$ et donc $f(x) = \lim_{n \rightarrow +\infty} \left(\frac{x^{2n} + 1}{x^{2n} - 1} \right) = -1$. Soit $x \in \mathbb{R}$ tel que $|x| > 1$, alors $\lim_{n \rightarrow +\infty} (|x^{2n}|) = +\infty$ donc $f(x) = \lim_{n \rightarrow +\infty} \left(\frac{x^{2n} + 1}{x^{2n} - 1} \right) = 1$. Nous avons donc : $f(x) = -1$ si $x \in]-1; +1[$, $f(x) = 0$ si $x = -1$ ou $x = 1$ et $f(x) = 1$ si $x \in]-\infty; -1[\cup]+1; +\infty[$. f est donc continue en tout point de $\mathbb{R}$ différent de -1 et 1 . En -1 et 1 , f n'est pas continue.

36.6 La fonction f est définie sur $\mathbb{R} - \{-1\}$. De plus, la fonction f est continue sur $\mathbb{R} - \{-1\}$, comme la composée et le produit de fonctions continues. Nous avons :

$$\lim_{x \rightarrow -1} f(x) = \lim_{u \rightarrow 0} \frac{\sin(u)}{u} (u \ln |u|) = 0.$$

Nous pouvons donc prolonger f par continuité en $x = -1$ et par conséquent sur $\mathbb{R}$.

36.7

1. La fonction f est définie et continue sur $\mathbb{R}^*$. Le seul problème est donc en $x = 0$. Nous savons que $\lim_{x \rightarrow 0} \sin(x) = 0$ et pour tout $x \neq 0$, $\left| \sin\left(\frac{1}{x}\right) \right| \leq 1$. Donc $\lim_{x \rightarrow 0} f(x) = 0$. La fonction f se prolonge donc par continuité sur $\mathbb{R}$, en posant $f(0) = 0$.

2. La fonction f est définie et continue sur $\mathbb{R}^*$. Le seul problème est donc en $x = 0$. Nous savons que $\lim_{x \rightarrow 0} \cos(x) = 1$ et pour tout $x \neq 0$, $\cos\left(\frac{1}{x}\right)$ oscille entre -1 et 1 . Donc la fonction f n'a pas de limite en $x = 0$. La fonction f ne se prolonge donc pas par continuité sur $\mathbb{R}$.

Fiche 37

37.1 La fonction f est dérivable sur $\mathbb{R}$ car f est dérivable comme quotient de deux fonctions polynomiales dont celle qui est au dénominateur ne s'annule jamais sur $\mathbb{R}$:

$$\begin{aligned} f'(x) &= \frac{2x(2x^4 + 8) - (x^2 + 3)8x^3}{(2x^4 + 8)^2} \\ &= \frac{-4x^5 - 24x^3 + 16x}{(2x^4 + 8)^2}. \end{aligned}$$

37.2 La fonction f est trois fois dérivable sur $\mathbb{R}$. Posons : $h(x) = \cos(7x + 2)$ et $g(x) = e^{-x^2}$. Nous avons :

$$f^{(3)}(x) = \sum_{k=0}^3 \binom{3}{k} h^{(k)}(x) g^{(3-k)}(x).$$

D'une part, nous avons $h^{(1)}(x) = -7 \sin(7x + 2)$, $h^{(2)}(x) = -49 \cos(7x + 2)$, $h^{(3)}(x) = 343 \sin(7x + 2)$. D'autre part, nous avons $g^{(1)}(x) = -2xe^{-x^2}$, $g^{(2)}(x) = -2e^{-x^2} + 4x^2e^{-x^2}$, $g^{(3)}(x) = 12xe^{-x^2} - 8x^3e^{-x^2}$. Ainsi, nous avons $f^{(3)}(x) = e^{-x^2}(\cos(7x+2)(306x-8x^3) + \sin(7x+2) \times (385-84x^2))$.

37.3 La fonction f est dérivable sur $\mathbb{R}$ comme fonction rationnelle de deux fonctions polynomiales dont le dénominateur ne s'annule jamais sur $\mathbb{R}$. Donc pour tout $x \in \mathbb{R}$, $f'(x) = \frac{3x^2 + 8x}{x^4 + x^2 + 1} - \frac{(x^3 + 4x^2 - 11)(4x^3 + 2x)}{(x^4 + x^2 + 1)^2}$.

37.4 La fonction f est dérivable sur $]2; +\infty[$ comme fonction rationnelle de deux fonctions dérivables sur $]2; +\infty[$. Donc pour tout $x \in]2; +\infty[$, $f'(x) = \frac{6x}{(x^2 - 4)\sqrt{x^3 + 4}} - \frac{9x^2 \ln(x^2 - 4)}{2(x^3 + 4)^{3/2}}$.

37.5 En appliquant la formule de Leibniz, nous remarquons que :

$$\text{si } k \geq 4, \text{ alors } \frac{d^k}{dx^k}(x^3 + x^2 + 1) = 0.$$

Ainsi, après calculs, nous obtenons :

$$\forall x \in \mathbb{R}, \quad f^{(n)}(x) = (-1)^n (x^3 - (3n-1)x^2 + (3n^2 - 5n)x - (n^3 - 4n^2 + 3n - 1)) \exp(-x).$$

37.6 Nous décomposons d'abord en éléments simples et nous obtenons

$$\forall x \in]-1; +1[, \quad f(x) = \frac{5/2}{(x-1)^2} - \frac{1/4}{x-1} + \frac{1/4}{x+1}.$$

Ainsi, nous avons :

$$\forall x \in]-1; +1[, \quad f^{(n)}(x) = \frac{5}{2} \frac{(-1)^{n+1}(n+1)!}{(x-1)^{n+2}}$$

$$- \frac{1}{4} \frac{(-1)^n n!}{(x-1)^{n+1}} + \frac{1}{4} \frac{(-1)^n n!}{(x+1)^{n+1}}.$$

Fiche 38

38.1 La fonction f est continue sur $[0; 1]$, dérivable sur $]0; 1[$ et $f(0) = f(1) = 4$. Donc d'après le théorème de Rolle, il existe $c \in]0; 1[$ tel que $f'(c) = 0$.

38.2 Soient $x < y \in [0; \frac{1}{2}]$. La fonction f est convexe donc

$$\begin{aligned} \frac{f(y) - f(x)}{y - x} &\leq \frac{f(1 - y) - f(y)}{1 - y - y} \\ &\leq \frac{f(1 - x) - f(1 - y)}{1 - x - (1 - y)} = \frac{f(1 - x) - f(1 - y)}{y - x} \end{aligned}$$

donc $f(y) - f(x) \leq f(1 - y) - f(1 - x)$ car $y - x > 0$, donc $\phi(y) \leq \phi(x)$. La fonction ϕ est décroissante sur $[0; \frac{1}{2}]$.

38.4 La fonction réelle h de l'indication vérifie $h(a) = h(b) = 0$, donc par le théorème de Rolle, il existe $c \in]a; b[$ tel que $h'(c) = 0$. Ce qui donne le résultat.

38.5 La fonction f est deux fois dérivable comme produit de deux fonctions dérivables sur $\mathbb{R}$. Nous avons : $\forall x \in \mathbb{R}, f'(x) = 2xe^x + (x^2 + 4)e^x$ puis $f''(x) = e^x((x + 2)^2 + 2)$. Ainsi $\forall x \in \mathbb{R}, f''(x) \geq 0$. Donc la fonction f est convexe sur $\mathbb{R}$.

Fiche 39

39.1

1. Pour tout $x \in \mathbb{R}$, nous avons :

$$f'(x) = -2 \sin(2x) \sin(3x) + 3 \cos(2x) \cos(3x) + 2x \tan(8x^2) + 16x^3(1 + \tan(8x^2)^2).$$

2. Pour tout $x \in \mathbb{R}$, nous avons :

$$f'(x) = 15(3x + 4)^4(x^2 - \cos(x) + x)^3 + 3(3x + 4)^5(x^2 - \cos(x) + x)^2(2x + \sin(x) + 1).$$

3. Pour tout $x \in \mathbb{R}$, nous avons :

$$\begin{aligned} f''(x) &= \frac{1}{1 + \left(\frac{x^3 + \cos x}{x^2 + 3x + 1}\right)^2} \\ &\times \left(\frac{3x^2 - \sin x}{x^2 + 3x + 1} - \frac{(x^3 + \cos x)(2x + 3)}{(x^2 + 3x + 1)^2}\right). \end{aligned}$$

39.2 $\arcsin\left(\frac{5}{6}\right) + \arcsin\left(\frac{\sqrt{11}}{6}\right) = \frac{\pi}{2}$.

39.3 L'ensemble des solutions de l'équation est l'intervalle $[-1; +1]$.

39.4

1. D'abord nous montrons que, pour tout $x > 0$, la dérivée de $\arctan(x) + \arctan\left(\frac{1}{x}\right)$ est

ulle. Puis nous en déduisons que $\arctan(x) + \arctan\left(\frac{1}{x}\right)$ est constante. Par exemple, elle est égale à ce qu'elle vaut en $x = 1$, donc $\frac{\pi}{2}$.

2. Même chose, sauf qu'il faut prendre $x = -1$. Ou bien, utiliser la question 1. avec $-x > 0$, et utiliser le fait que la fonction $\arctan$ est impaire.

39.5 Posons $x = \tan(\theta)$ avec $\theta \in \left]-\frac{\pi}{2}; \frac{\pi}{2}\right[$. Alors nous avons $\sin(\arctan(x)) = \sin(\theta) = \tan(\theta) \cos(\theta) = x \cos(\theta) = x \frac{1}{\sqrt{1+x^2}}$ car $1 + \tan^2(\theta) = \frac{1}{\cos^2(\theta)}$. Donc nous écrivons $\cos \theta = \pm \frac{1}{\sqrt{1 + \tan^2(\theta)}}$. Comme, nous avons choisi $\theta \in \left]-\frac{\pi}{2}; \frac{\pi}{2}\right[$, alors $\cos \theta > 0$. Par conséquent, nous avons $\cos \theta = \frac{1}{\sqrt{1 + \tan^2(\theta)}}$ et l'égalité est bien démontrée.

39.6

1. Pour $x \neq \frac{\pi}{2}$, nous avons $-2 \cotan(2x) + \cotan(x) = -2 \frac{\cos(2x)}{\sin(2x)} + \frac{\cos(x)}{\sin(x)} = \frac{-2 \cos(2x)}{2 \sin(x) \cos(x)} + \frac{\cos(x)}{\sin(x)} = \frac{-\cos(2x)}{\sin(x) \cos(x)} + \frac{\sin^2(x)}{\sin(x) \cos(x)} = \frac{\sin(x)}{\cos(x)} = \tan(x)$, en utilisant les formules $\cos(2x) = \cos^2(x) - \sin^2(x)$ et $\sin(2x) = 2 \sin(x) \cos(x)$.

2. Après calculs, nous obtenons, pour $n \geq 1$, $S_n = -2^{n+1} \cotan(2^{n+1}x) + \cotan(x)$.

39.7

1. Pour $x \neq \frac{\pi}{2}$, nous avons $1 + \frac{1}{\cos(2x)} = \frac{2 \cos^2 x}{\cos(2x)}$ et de même $\frac{\tan(2x)}{\tan(x)} = \frac{2 \cos^2(x)}{\cos(2x)}$ en utilisant que $\cos(2x) = 2 \cos^2(x) - 1$ et $\sin(2x) = 2 \sin(x) \cos(x)$.

2. Après calculs, nous obtenons, pour $n \geq 1$, $S_n = \frac{\tan(2^n x)}{\tan\left(\frac{x}{2}\right)}$.

$$\begin{aligned} \text{39.8} \quad \cos(2a) &= \frac{1 - \sqrt{5}}{4}, \\ \cos(4a) &= -\frac{1 + \sqrt{5}}{4}, \quad a = \frac{3\pi}{10}. \end{aligned}$$

Fiche 40

40.1 L'équation peut s'écrire, pour tout x réel, sous la forme :

$$(\operatorname{ch}(x) + \operatorname{sh}(x))^4 - 3 \operatorname{ch}(x) \operatorname{sh}(x) (\operatorname{ch}(x) + \operatorname{sh}(x))^2 = 0.$$

Après simplifications, nous avons, pour tout x réel :

$$e^{4x} - \frac{3}{4} (e^x + e^{-x}) (e^x - e^{-x}) e^{2x} = 0.$$

Après développements nous avons, pour tout x réel :

$$e^{4x} - \frac{3}{4} (e^{4x} - 1) = 0 \Leftrightarrow e^{4x} + 3 = 0.$$

Or, pour tout $x \in \mathbb{R}$, $e^{4x} + 3 > 0$. Donc l'ensemble des solutions est l'ensemble vide.

40.2 L'équation est équivalente à, pour tout x réel :

$$\operatorname{ch}(x) = \frac{1}{\operatorname{sh}(x)} \Leftrightarrow \operatorname{sh}(2x) = 2.$$

Nous en déduisons que l'ensemble des solutions est égal à l'unique valeur $\frac{\operatorname{argsh}(2)}{2}$.

40.3 Pour tout $x \in \mathbb{R}$, $f(x) = 2 \arctan(\operatorname{th}(x)) - \arctan(\operatorname{sh}(2x))$. La fonction f est dérivable sur $\mathbb{R}$. Par conséquent, pour tout $x \in \mathbb{R}$, nous avons

$$f'(x) = 2 \frac{1 - \operatorname{th}^2(x)}{1 + \operatorname{th}^2(x)} - \frac{2 \operatorname{ch}(2x)}{1 + (\operatorname{sh}(2x))^2}.$$

Puis

$$\begin{aligned} \frac{2 \operatorname{ch}(2x)}{1 + (\operatorname{sh}(2x))^2} &= \frac{2 \operatorname{ch}(2x)}{(\operatorname{ch}(2x))^2} = \frac{2}{\operatorname{ch}(2x)} = \\ \frac{2}{\operatorname{ch}^2(x) + \operatorname{sh}^2(x)} &= \frac{2/\operatorname{ch}^2(x)}{1 + \operatorname{th}^2(x)} = \frac{2(1 - \operatorname{th}^2(x))}{1 + \operatorname{th}^2(x)}. \end{aligned}$$

Ainsi, pour tout $n \in \mathbb{N}$, $f'(x) = 0$. Ainsi nous en déduisons que la fonction f est constante sur $\mathbb{R}$. Pour calculer cette constante, nous calculons par exemple $f(0)$ et nous trouvons que $f(0) = 0$. Ainsi, pour tout $x \in \mathbb{R}$, nous avons $f(x) = 0$. Donc, l'ensemble des solutions est $\mathbb{R}$.

Fiche 41

41.1

$$1. \quad f(x) = 1 + \frac{x}{2} + \frac{1 \cdot 3}{2 \cdot 4} x^2 + \dots + \frac{1 \cdot 3 \dots (2n-1)x^n}{2 \cdot 4 \dots (2n)} + o(x^n).$$

$$2. \quad f(x) = \ln(2+x) = \ln 2 + \ln\left(1 + \frac{x}{2}\right) = \ln 2 + \frac{x}{2} - \frac{x^2}{8} + \frac{x^3}{3 \cdot 2^3} + \dots + \frac{(-1)^{n-1} x^n}{n \cdot 2^n} + o(x^n).$$

$$3. \quad \frac{f(x)}{\sin(3x) - \sin(x)} = \frac{\sin(x) \cos(2x)}{\sin(3x) - \sin(x)} = \frac{\sin(x) \cos(2x)}{2} = x - \frac{13x^3}{6} + \frac{121x^5}{5!} + o(x^6).$$

$$4. \quad f(x) = x + \frac{x^2}{2} + \frac{x^3}{3} + o(x^3).$$

41.2

1. En appliquant l'inégalité de Taylor-Lagrange à $f(x) = \cos(x)$ pour $n = 4$, $x = a$, $x_0 = 0$ et $M = 1$, nous trouvons le résultat souhaité.

2. En prenant $a = \frac{1}{2}$, nous obtenons le résultat demandé.

41.3

1. En appliquant l'inégalité de Taylor-Lagrange à $f(x) = \ln(1+x)$ pour $n = 2$, $x \geq 0$, $x_0 = 0$ et $M = 2$, nous trouvons le résultat souhaité.

2. En appliquant le résultat précédent avec $x = 0,003$, nous trouvons $\ln(1,003) \approx 0,00299551$.

41.4

L'inégalité de Taylor-Lagrange pour $f(x) = \ln(1+x)$, à l'ordre n , $x > 0$, en $x_0 = 0$ et $M = n!$ donne $\left| \ln(1+x) - x + \frac{x^2}{2} - \frac{x^3}{3} + \dots + (-1)^{n-1} \frac{x^n}{n} \right| \leq \frac{x^{n+1}}{n+1}$ pour tout $x > 0$. Donc pour $x = 1$, $|\ln 2 - u_n| \leq \frac{1}{n+1} \rightarrow 0$ lorsque $n \rightarrow +\infty$.

41.5

Nous appliquons deux fois l'égalité de Taylor-Young à l'ordre deux et nous obtenons

$$\begin{aligned} f(a+h) &= f(a) + hf'(a) + \frac{h^2}{2} f''(a) + o_1(h^2) \\ \text{et} \\ f(a-h) &= f(a) - hf'(a) + \frac{h^2}{2} f''(a) + o_2(h^2). \end{aligned}$$

Ainsi, nous en déduisons que

$$\lim_{h \rightarrow 0} \frac{f(a+h) + f(a-h) - 2f(a)}{h^2} = f''(a).$$

Fiche 42

$$42.1 \quad DL(0; 2n+1; \exp) = \sum_{k=0}^{2n+1} \frac{x^k}{k!} + o(x^{2n+1}).$$

Pour tout $x \in \mathbb{R}^*$:

$$\begin{aligned} f_n(x) &= \frac{1}{x^{n+1}} \left(\exp(x) - \sum_{k=0}^n \frac{x^k}{k!} \right) \\ &= \frac{1}{x^{n+1}} \left(\sum_{k=n+1}^{2n+1} \frac{x^k}{k!} + o(x^{2n+1}) \right) \\ &= \sum_{k=0}^n \frac{x^k}{(k+n+1)!} + o(x^n). \end{aligned}$$

Par conséquent, $\lim_{x \rightarrow 0} f_n(x) = \frac{1}{(n+1)!}$ et f_n se prolonge par continuité en 0.

42.2 Nous savons que $DL(0; 1; \sin) = x + o(x^2)$. La fonction arc sinus est dérivable en 0 et admet donc un développement limité à l'ordre 1 en 0 : $\arcsin y = a_0 + a_1 y + o(y)$. Or $\arcsin(0) = 0$ donc $a_0 = 0$. Au voisinage de 0, $\sin(\arcsin y) = \sin(a_1 y + o(y)) = a_1 y + o(y) = y$ car $\sin \circ \arcsin = id$. Ainsi $a_1 = 1$.

Par conséquent $\frac{\arcsin y}{y}$ tend vers 1 lorsque y tend vers 0 et ainsi $\left(\frac{\arcsin y}{y} \right)^2 = 1 + o(1)$ en $y = 0$. Nous en déduisons le développement limité en 0 demandé :

$$\arcsin(y)^2 = y^2 + o(y^2).$$

42.3 $DL(3; 1; f)$. $\forall x \in \mathbb{R}_+^*$, $f(x) = e^{\left(1 + \frac{1}{x}\right) \ln x}$. Posons $x = 1 + h$. Pour tout $h \in]-1; +\infty[$:

$$\begin{aligned} f(h) &= e^{(1 + \frac{1}{1+h}) \ln(1+h)} \\ &= e^{(1 + 1 - h + h^2 - h^3 + o(h^3)) \left(h - \frac{h^2}{2} + \frac{h^3}{3} + o(h^3) \right)} \\ &= e^{2h - h^2 + \frac{2}{3}h^3 - h^2 + \frac{1}{2}h^3 + h^3 + o(h^3)} \\ &= e^{2h - 2h^2 + \frac{13}{6}h^3 + o(h^3)} \\ &= 1 + \left(2h - 2h^2 + \frac{13}{6}h^3 \right) \\ &\quad + \frac{(2h - 2h^2 + \frac{13}{6}h^3)^2}{2} + \frac{(2h - 2h^2 + \frac{13}{6}h^3)^3}{6} \\ &\quad + o(h^3) \\ &= 1 + 2h - 2h^2 + 2h^2 + \frac{13}{6}h^3 - 4h^3 + \frac{8}{6}h^3 + o(h^3) \\ &= 1 + 2h - \frac{1}{2}h^3 + o(h^3). \end{aligned}$$

$$\begin{aligned} 42.4 \quad f_1(\lambda) &= \frac{8\pi kT}{\lambda^4} \text{ et } f_2(\lambda) = \frac{8\pi h c \lambda^{-5}}{e^{hc/\lambda kT} - 1}. \\ \text{Posons } x &= 1/\lambda, x \text{ est voisin de } 0. \text{ Nous} \\ \text{avons } f_2(\lambda) &= \frac{8\pi h c x^5}{\frac{hc x}{kT} + o(x)} = \frac{x}{kT} (1 + o(1)) = \\ 8\pi kT x^4 (1 - o(1)) &= \frac{8\pi kT}{\lambda^4} + o\left(\frac{1}{\lambda^4}\right) = f_1(\lambda) + \\ o\left(\frac{1}{\lambda^4}\right). \end{aligned}$$

Fiche 43

43.1

1. Domaine de définition : $\mathbb{R} \setminus \{-2; 2\}$.

f n'est ni paire, ni impaire, ni périodique.
 f est strictement décroissante sur les intervalles $] -\infty; -4 - 2\sqrt{3}]$, $] -4 + 2\sqrt{3}; 2[$ et $]2; +\infty[$.
 f est strictement croissante sur les intervalles $[-4 - 2\sqrt{3}; -2[$ et $] -2; -4 + 2\sqrt{3}]$.

La droite $x = -2$ est asymptote parallèle à (Oy) .

La droite $x = 2$ est asymptote parallèle à (Oy) .

La droite $y = 1$ est asymptote parallèle à l'axe des abscisses au voisinage de $-\infty$ et $+\infty$.

2. Domaine de définition : $\mathbb{R}$.

f est impaire et périodique de période π .

La courbe est symétrique par rapport au point $(0; 0)$ (car f est impaire) et au point $(\frac{\pi}{2}; 0)$.

Par conséquent, nous restreignons l'intervalle d'étude à $[0; \frac{\pi}{2}]$.

f est strictement croissante sur les intervalles $[0; 0, 28]$ et $[1, 1; \frac{\pi}{2}]$ et strictement décroissante sur l'intervalle $[0, 28; 1, 1]$.

3. Domaine de définition : $\mathbb{R}$.

Il n'y a pas de symétrie.

f est strictement croissante sur l'intervalle $] -\infty; 1/\ln(2)]$ et strictement décroissante sur l'intervalle $[1/\ln(2); +\infty[$.

La droite $y = 0$ est asymptote en $+\infty$.

43.2

1. Si f est paire, elle ne peut pas être bijective, car elle n'est pas injective !

Soit f une fonction impaire, bijective et f^{-1} sa réciproque. Pour tout $x \in \mathbb{R}$, $f(x) = -f(-x)$.

Soit $y \in \mathbb{R}$ tel que $y = f(x)$, alors $-y = -f(x) = f(-x)$, donc $f^{-1}(-y) = -x = -f^{-1}(y)$. Donc f^{-1} est impaire.

2. Les fonctions $\tan$ et $\tanh$ sont impaires, donc leurs réciproques $\arctan$ et argth le sont aussi.

Fiche 44

44.1 Lorsque $n \rightarrow +\infty$, $v_n = \frac{1}{n} \sum_{k=1}^n \frac{1}{1 + \left(\frac{k}{n}\right)^2} \rightarrow \int_0^1 f(x) dx$ avec $f(x) = \frac{1}{1+x^2}$. D'où $\lim_{n \rightarrow +\infty} v_n = [\arctan x]_0^1 = \frac{\pi}{4}$.

44.2 La fonction définie par $f(x) = x$ si $x < 1$ et $f(x) = -x$ si $x \geq 1$ est continue sur $[0; 1[\cup]1; 2]$. Mais $\lim_{x \rightarrow 1^-} f(x) = 1$ et $\lim_{x \rightarrow 1^+} f(x) = -1$. La fonction f n'est pas continue en 1. Elle est continue par morceaux sur $[0; 2]$.

Fiche 45

45.1 Pour $x \in]0; \pi[$, une primitive de $f(x)$ est égale à $\ln\left(\tan\left(\frac{x}{2}\right)\right) + k$, $k \in \mathbb{R}$.

45.2 Pour $x > 0$, une primitive de $f(x)$ est égale à $\frac{\sqrt{x^{10}+1}}{5} - \frac{1}{10} \ln \left| \frac{1 + \sqrt{x^{10}+1}}{1 - \sqrt{x^{10}+1}} \right| + k$, $k \in \mathbb{R}$.

45.3

1. Pour $x \in \mathbb{R}$, une primitive de $f(x)$ est égale à $\frac{1}{16} \ln(x^2 + 2x + 2) + \frac{1}{8} \arctan(x+1) - \frac{1}{16} \ln(x^2 - 2x + 2) + \frac{1}{8} \arctan(x-1) + k$, $k \in \mathbb{R}$.

2. Pour $x \neq 2$ et pour $x \neq \frac{1}{2}$, une primitive de $f(x)$ est égale à $3 \ln(x+1) - \frac{1}{2} \ln(2x-1) + k$, $k \in \mathbb{R}$.

3. Pour $x \neq 0$, une primitive de $f(x)$ est égale à $\frac{1}{2} \ln(x^2 + 4) - \frac{1}{2} \arctan\left(\frac{x}{2}\right) + \ln(x) + k$, $k \in \mathbb{R}$.

45.4 Pour $x > -1$, une primitive de $f(x)$ est égale à $x \left(1 + \frac{1}{2}x\right) \arctan\left(\sqrt{\frac{x+3}{x+1}}\right) +$

$\frac{1}{4} \sqrt{3+4x+x^2} - \frac{1}{2} \ln(x+2 + \sqrt{3+4x+x^2}) + k$, $k \in \mathbb{R}$.

45.5 Pour $x \in]-1; +1[$, une primitive de $f(x)$ est égale à $\frac{3}{2} \arcsin(x) - \sqrt{1-x^2} - \frac{1}{2}x\sqrt{1-x^2} + k$, $k \in \mathbb{R}$.

45.6 Pour $x \in \mathbb{R}$, une primitive de $f(x)$ est égale à $-\frac{2}{\tanh\left(\frac{x}{2}\right) + 3} + k$, $k \in \mathbb{R}$.

45.7 Pour $x \neq 0 + 2k\pi$, avec $k \in \mathbb{Z}$, une primitive de $f(x)$ est égale à $3 \ln \left| \frac{1 + 2 \cos\left(\frac{x}{6}\right)}{1 - 2 \cos\left(\frac{x}{6}\right)} \right| + k$, $k \in \mathbb{R}$.

Fiche 46

46.1

1. Il s'agit d'une série alternée, telle que $|u_n|$ tend vers 0 en décroissant, donc la série est convergente.

2. Il s'agit d'une série alternée, telle que $|u_n|$ tend vers 0 en décroissant, donc la série est convergente.

Fiche 47

47.1

1. Une application de la règle de d'Alembert permet de montrer que la série converge :

$$\frac{u_{n+1}}{u_n} = e^{2^n \ln 3 - 3^n 2 \ln 2} \rightarrow_{n \rightarrow +\infty} 0.$$

2. $\ln\left(\frac{n^2 + n + 1}{n^2 + n - 1}\right)$ est équivalent à $\frac{2}{n^2}$ donc la série converge.

3. $0 \leq n^2 \left(\sqrt{n} + \sqrt[3]{n}\right)^{-\sqrt{n}} \leq n^2 2^{-\sqrt{n}} = e^{2 \ln n - \sqrt{n} \ln 2} \rightarrow_{n \rightarrow +\infty} 0$. Par conséquent, le terme de la série est négligeable devant $\frac{1}{n^2}$ qui est le terme général d'une série absolument convergente.

4. Terme général équivalent à $\frac{1}{n}$ lorsque n tend vers $+\infty$. Donc la série diverge.

5. $\left| \frac{\sin n}{n^2} \right| \leq \frac{1}{n^2}$ qui est le terme général d'une série absolument convergente. Puis $(-1)^n n^{-\frac{3}{2}}$ est le terme général d'une série qui vérifie les hypothèses du théorème spécial des séries alternées : elle converge également. Ainsi la série de terme général $\frac{(-1)^n \sqrt{n} + \sin n}{n^2}$ est convergente puisqu'elle est la somme de deux séries convergentes.

6. $\left| \left(1 - \frac{n}{\ln n}\right)^{-n} \right| = \left(\frac{n}{\ln n} - 1\right)^{-n} \leq 2^{-n}$ lorsque n tend vers $+\infty$. La série est absolument convergente donc convergente.

47.2 En appliquant la Proposition 47.8 à la fonction $[1; +\infty[\ni x \mapsto \frac{1}{x} \in \mathbb{R}$, nous obtenons pour tout $n \in \mathbb{N}^*$:

$$\int_1^{n+1} \frac{1}{x} dx \leq \sum_{i=1}^n \frac{1}{k} \text{ et } \sum_{i=2}^n \frac{1}{k} \leq \int_1^n \frac{1}{x} dx.$$

D'où, pour tout $n \in \mathbb{N}^*$, $\ln(n+1) \leq \sum_{i=1}^n \frac{1}{k} \leq 1 + \ln n$. Nous en déduisons l'équivalent demandé puisque $\ln(n+1)$ est équivalent à $\ln n$ lorsque $n \rightarrow +\infty$.

47.3

1. Les égalités suivantes, vraies pour $n \geq 1$, aboutissent à l'équivalent demandé lorsque $n \rightarrow +\infty$:

$$\begin{aligned} v_n &= u_{n+1} - u_n \\ &= \sum_{k=1}^{n+1} \frac{1}{k} - \ln(n+1) - \sum_{k=1}^n \frac{1}{k} + \ln n \\ &= \frac{1}{n+1} - \ln\left(1 + \frac{1}{n}\right) \\ &= \frac{1}{n} \left(\frac{1}{1 + \frac{1}{n}} \right) - \ln\left(1 + \frac{1}{n}\right) \\ &= \frac{1}{n} \left(1 - \frac{1}{n} + \frac{1}{n^2} + o\left(\frac{1}{n^2}\right) \right) \\ &\quad - \left(\frac{1}{n} - \frac{1}{2n^2} + o\left(\frac{1}{n^2}\right) \right) \text{ lorsque } n \rightarrow +\infty \\ &= -\frac{1}{2n^2} + o\left(\frac{1}{n^2}\right). \end{aligned} \quad (47.1)$$

2. Le terme général v_n de la série $\sum_{n \geq 1} v_n$ est équivalent à celui $w_n = -\frac{1}{2n^2}$ de la série $\sum_{n \geq 1} w_n$ absolument convergente. Par conséquent la série $\sum_{n \geq 1} v_n$ converge.

3. Remarquons que $\sum_{n=1}^k v_n = u_{k+1} - u_1$ ainsi la convergence de la série $\sum_{n \geq 1} v_n$ implique celle de la suite $(u_n)_{n \in \mathbb{N}^*}$. La limite de cette suite est appelée constante d'Euler et est généralement notée γ .

47.4

1. Notons $\beta = \frac{\alpha+1}{2}$ et, pour tout $n \in \mathbb{N}^*$, $v_n = \frac{1}{n^\beta}$.

La série $\sum v_n$ est une série de Riemann qui converge puisque $\beta > 1$ car nous avons $\alpha > 1$. Puis, nous avons lorsque $n \rightarrow +\infty$:

$$\frac{v_{n+1}}{v_n} = \left(1 + \frac{1}{n}\right)^{-\beta} = 1 - \frac{\beta}{n} + o\left(\frac{1}{n}\right).$$

Notons $\gamma = \frac{\alpha+\beta}{2}$ de sorte que $1 < \beta < \gamma < \alpha$. Il existe $N \in \mathbb{N}$ tel que :

$$\forall n \geq N, \frac{u_{n+1}}{u_n} \leq 1 - \frac{\gamma}{n} \leq \frac{v_{n+1}}{v_n}.$$

Par conséquent,

$$\frac{u_n}{u_{n-1}} \leq \frac{v_n}{v_{n-1}}, \frac{u_{n-1}}{u_{n-2}} \leq \frac{v_{n-1}}{v_{n-2}}, \dots, \frac{u_{N+1}}{u_N} \leq \frac{v_{N+1}}{v_N}$$

d'où par multiplication $u_n \leq \frac{u_N}{v_N} v_n$.

Le critère de comparaison, nous permet d'affirmer que la convergence de $\sum v_n$ implique celle de $\sum u_n$.

2. Si $\alpha \leq 0$, à partir d'un certain rang nous avons $u_{n+1} \geq u_n$ et $u_n > 0$ donc $\sum_n u_n$ diverge. Supposons maintenant que $\alpha > 0$. Il suffit alors de raisonner comme à la première question pour conclure.

Fiche 48

48.1 Pour tout $x \in [0; 1]$, la suite de fonctions (f_n) converge simplement sur $[0; 1]$ vers la fonction f définie par $f(x) = \frac{e^{-x}}{2}$.

La suite de fonctions (f_n) converge aussi uniformément sur $[0; 1]$ vers la fonction f , car

pour tout $x \in [0; 1]$, lorsque $n \rightarrow +\infty$

$$\left| f_n(x) - \frac{e^{-x}}{2} \right| = \left| \frac{3x^3 - e^{-x}}{2n+2} \right| \leq \frac{4}{2n+2} \rightarrow 0.$$

Fiche 50

50.1 Le rayon de convergence est $R = \frac{1}{3}$.
L'intervalle de convergence est $] -\frac{1}{3}; +\frac{1}{3}[$.

50.2 Le rayon de convergence est $R = 0$.

50.3 L'intervalle de convergence est $] -5; 1[$.

50.4 Le rayon de convergence est $R = 1$.
L'intervalle de convergence est $] -1; 1[$.

Fiche 51

51.1 L'intervalle de convergence est égal à $] -1; +1[$.

$$\frac{1}{1+x^2} = \sum_{n=0}^{+\infty} (-1)^n x^{2n}.$$

51.2 L'intervalle de convergence est égal à $] -2; +2[$.

$$\frac{1}{2+x} = \sum_{n=0}^{+\infty} \frac{(-1)^n}{2^{n+1}} x^n.$$

51.3 L'intervalle de convergence est égal à $\mathbb{R}$.

$$\cos(2x) = \sum_{n=0}^{+\infty} (-1)^n \frac{2^{2n}}{(2n)!} x^{2n}.$$

51.4 L'intervalle de convergence est égal à $\mathbb{R}$.

$$\operatorname{sh}\left(\frac{x}{2}\right) = \sum_{n=0}^{+\infty} \frac{1}{2^{2n+1}(2n+1)!} x^{2n+1}.$$

51.5 L'intervalle de convergence est égal à $] -1; +1[$.

$$\ln(1+x) = \sum_{n=1}^{+\infty} (-1)^{n+1} \frac{x^n}{n}.$$

51.6 L'intervalle de convergence est égal à $] -2; +2[$.

$$\begin{aligned} \frac{x^3}{2+x} &= \frac{x^3}{2} \frac{1}{1+\frac{x}{2}} = \frac{x^3}{2} \sum_{n=0}^{+\infty} (-1)^n \frac{x^n}{2^n} \\ &= \sum_{n=0}^{+\infty} (-1)^n \frac{x^{n+3}}{2^{n+1}}. \end{aligned}$$

51.7 L'intervalle de convergence est égal à $] -2; +2[$.

$$\ln(x^2 - 5x + 6) = \ln 6 - \sum_{n=1}^{+\infty} \left(\frac{1}{3^n} + \frac{1}{2^n} \right) \frac{x^n}{n}.$$

51.8 L'intervalle de convergence est égal à $] -\frac{1}{2}; +\frac{1}{2}[$.

$$\ln(1+x-2x^2) = \sum_{n=1}^{+\infty} \frac{(-1)^{n+1} 2^n - 1}{n} x^n.$$

51.9 L'intervalle de convergence est égal à $] -1; +1[$.

$$(\arcsin x)^2 = 2 \sum_{n=0}^{+\infty} \frac{(2^n n!)^2}{(2n+2)!} x^{2n+2}.$$

51.10 L'intervalle de convergence est égal à $] -1; +1[$.

$$\sqrt{x + \sqrt{x^2 + 1}} = \sum_{n=0}^{+\infty} a_n x^n$$

$$\text{où } a_{2k} = \frac{1}{4^k (2k)!} \prod_{j=1}^{k-1} 1 - 16j^2 \text{ et } a_{2k+1} = \frac{1}{2 \cdot 4^k (2k+1)!} \prod_{i=1}^k 1 - 4(2i-1)^2.$$

51.11 Si $a_0 = a_1 = a_2 = 0$, alors, pour tout $n \in \mathbb{N}$, $a_n = 0$. Le rayon de convergence est infini.

Sinon, l'intervalle de convergence est égal à $] -1; +1[$ et

$$\sum_{n=0}^{+\infty} a_n x^n = \frac{a_0 + (a_1 - a_0)x + (a_2 - a_1 - a_0)x^2}{(1-x)^2(1+x)}.$$

51.12

$$\sum_{n=0}^{+\infty} \frac{x^{3n}}{(3n)!} = \frac{1}{3} \left(e^x + e^{-\frac{x}{2}} \cdot 2 \cos\left(\frac{\sqrt{3}}{2}x\right) \right).$$

51.13 L'intervalle de convergence est égal à $] - 1; +1[$ et

$$\sum_{n=0}^{+\infty} \frac{x^n}{n} \cos\left(\frac{2n\pi}{3}\right) = -\frac{1}{2} \ln(1 + x + x^2).$$

Fiche 53

53.1

1. Intégrale divergente.
2. Intégrale convergente.
3. Intégrale convergente.
4. Intégrale convergente.

53.2

1. Posons $h(x) = \frac{\sin^2 x}{x}$; $\frac{\sin^2 x}{\frac{x}{\sin x}} = \frac{\sin x}{\sqrt{x}} \rightarrow 0$

quand $x \rightarrow +\infty$. Ainsi $f = g + h$ avec $h = o(g)$ au voisinage de $+\infty$ ce qui montre que $f \sim g$ au voisinage de $+\infty$.

2. L'intégrale de g sur $[a; +\infty[$ converge alors que celle f diverge.

53.5

1. La fonction $f : x \in]0; \pi/2] \mapsto \ln(\sin x)$ est continue sur $]0; \pi/2]$ et a un signe constant sur cet intervalle car pour tout $x \in]0; \pi/2]$, $0 < \sin x \leq 1$ et ainsi $\ln(\sin x) \leq 0$.

Étude en 0 : $\ln(\sin x) \sim \ln x$ et $\ln$ est intégrable sur $]0; \pi/2]$ et y garde un signe constant sur $]0; \pi/2]$.

En conclusion, f est intégrable sur $]0; \pi/2]$.

Procédons au changement de variable $y = \frac{\pi}{2} - x$, avec $dy = -dx$, qui est licite car affine et est donc un $\mathcal{C}^1$ difféomorphisme de $]0; \pi/2]$ sur $]0; \pi/2[$:

$$J = \int_0^{\pi/2} \ln(\sin(x)) dx = - \int_{\pi/2}^0 \ln(\sin(\pi/2 - y)) dy = \int_0^{\pi/2} \ln(\cos(y)) dy.$$

$$2. \quad 2J = J + J = \int_0^{\pi/2} \ln(\sin(x)) dx + \int_0^{\pi/2} \ln(\cos(x)) dx = \int_0^{\pi/2} \ln(\sin(x) \cos(x)) dx = \int_0^{\pi/2} \ln\left(\frac{\sin(2x)}{2}\right) dx.$$

3. Commençons par remarquer que :

$$\int_{\pi/2}^{\pi} \ln(\sin x) dx = \int_{\pi/2}^0 \ln(\sin(\pi - z)) (-dz) =$$

$\int_0^{\pi/2} \ln(\sin z) dz = J$ avec $z = \pi - x$ et $dz = -dx$ qui est un changement de variable licite, car affine, de $[\pi/2; \pi[$ sur $]0; \pi/2]$.

Puis que : $J + J = \int_0^{\pi/2} \ln(\sin x) dx + \int_{\pi/2}^{\pi} \ln(\sin x) dx = \int_0^{\pi} \ln(\sin x) dx = \int_0^{\pi/2} \ln(\sin 2y) 2dy$ avec $x = 2y$ et $dx = 2dy$ qui est un changement de variable licite, car linéaire, de $]0; \pi[$ sur $]0; \pi/2[$.

$$2J = \int_0^{\pi/2} \ln\left(\frac{\sin 2x}{2}\right) dx = \int_0^{\pi/2} \ln(\sin 2x) dx - \int_0^{\pi/2} \ln 2 dx = \int_0^{\pi/2} \ln(\sin 2x) dx - \frac{\pi \ln 2}{2} = J - \frac{\pi \ln 2}{2}.$$

Par conséquent $J = -\frac{\pi \ln 2}{2}$.

Fiche 54

- 54.1 Introduisons la suite (f_n) , définie, pour tout $t \in [0; 1]$ et $n \in \mathbb{N}$, par $f_n(t) = \frac{1}{1 + t + t^2 + \dots + t^n}$.

Remarquons que, pour tout $n \in \mathbb{N}$, $f_n(1) = \frac{1}{n+1}$ et pour tout $t \in [0; 1[$ $f_n(t) = \frac{1-t}{1-t^{n+1}}$. Puis (f_n) converge simplement sur $[0; 1]$ vers la fonction qui à $t \in [0; 1] \mapsto 1-t \in \mathbb{R}$, fonction qui est continue sur $[0; 1]$.

Le théorème de convergence bornée, ou monotone car (f_n) décroît, permet d'affirmer que la suite (u_n) converge vers

$$\int_0^1 (1-t) dt = \frac{1}{2}.$$

Remarquons que la convergence de la suite (u_n) était assurée par le fait qu'elle est minorée par 0, puisque son terme général est l'intégrale d'une fonction positive et décroissante. En effet, nous avons $u_{n+1} - u_n = - \int_0^1 \frac{t^{n+1}}{(1+t+\dots+t^n)(1+t+\dots+t^n+t^{n+1})} dt \leq 0$.

- 54.2 Posons $f(t) = \frac{t}{\operatorname{sh} t}$ pour tout $t \in]0; +\infty[$. L'intégrale existe car f se prolonge par continuité en 0 et est positive au voisinage de $+\infty$ et équivalente à $2t^3 e^{-t}$ donc est un $o(t^{-2})$.

En écrivant $f(t) = \frac{2t}{e^t} \frac{1}{1 - e^{-2t}}$ et développant le terme $\frac{1}{1 - e^{-2t}}$ en série entière nous obtenons :

$$f(t) = 2t \sum_{n=0}^{+\infty} e^{-2(n+1)t}.$$

Le théorème d'intégration des séries de fonctions positives permet alors d'affirmer que :

$$\begin{aligned} \int_0^{+\infty} f(t) dt &= \int_0^{+\infty} 2t \sum_{n=0}^{+\infty} e^{-2(n+1)t} dt \\ &= 2 \sum_{n=0}^{+\infty} \int_0^{+\infty} t e^{-2(n+1)t} dt = 2 \sum_{n=0}^{+\infty} \frac{1}{(2k+1)^2}. \end{aligned}$$

Enfin utilisons l'égalité donnée dans l'énoncé :

$$\begin{aligned} \frac{\pi^2}{6} &= \sum_{n=0}^{+\infty} \frac{1}{n^2} = \sum_{k=0}^{+\infty} \frac{1}{(2k+1)^2} + \sum_{k=0}^{+\infty} \frac{1}{(2k)^2} \\ &= \frac{I}{2} + \frac{1}{4} \frac{\pi^2}{6}. \end{aligned}$$

Par conséquent, $I = \pi^2/4$.

54.3 Application directe du théorème de continuité sous le signe $\int$. En effet, soit ϕ définie sur $D =]-1; +\infty[\times]0; 1]$ par, si $t > 0$, $\phi(x, t) = \frac{\ln(1 + xt^2)}{t^2}$ et, si $t = 0$, $\phi(x, 0) = x$.

1. Pour tout $x \in]-1; +\infty[$, la fonction $t \mapsto \phi(x, t)$ est continue sur $]0; 1]$.

2. Pour tout $t \in]0; 1]$, la fonction $x \mapsto \phi(x, t)$ est continue sur $] -1; +\infty[$.

3. Soit $K = [a, b]$ un segment inclus dans $] -1; +\infty[$. Pour tout $(x, t) \in K \times]0; 1]$, $|\phi(x, t)| \leq \frac{\ln(1 + at^2)}{t^2}$ qui est continue, positive intégrable sur $]0; 1]$.

En conclusion, f est définie et continue sur tout segment de $] -1; +\infty[$, donc sur $] -1; +\infty[$ en entier.

Remarquons que, pour une fonction f définie sur $A \times K$, où A est une partie de $\mathbb{R}^m$ et K un segment de $\mathbb{R}$, la condition de domination dans le théorème de continuité est toujours vérifiée lorsque la fonction f est continue sur $A \times K$. La notion de continuité des fonctions de plusieurs variables est introduite à la Fiche 58.

Il en va de même pour le théorème de dérivation de Leibniz si f est de classe $\mathcal{C}^1$ sur $A \times K$.

La notion de classe $\mathcal{C}^1$ des fonctions de plusieurs variables est introduite à la Fiche 59.

54.4 f est définie sur $]0; \pi[$ en entier. Une application directe du théorème de dérivation sous le signe $\int$, la domination étant claire sur tout segment de $]0; \pi[$, permet d'affirmer que f est de classe $\mathcal{C}^1$ sur tout segment de $]0; \pi[$, donc sur $]0; \pi[$ en entier.

Puis un calcul permet de montrer que $f'(x) = \pi - x$ pour tout $x \in]0; \pi[$ et d'obtenir ainsi la valeur explicite de f :

$$\forall x \in]0; \pi[, f(x) = \frac{\pi^2}{6} - \frac{(x - \pi)^2}{2}.$$

54.6 Soient $f(x, t) = \frac{1 - \cos t}{\sqrt{\sin^2 t + x \cos^2 t}}$

et $F(x) = \int_0^{\frac{\pi}{2}} f(x, t) dt$.

1. Cette fonction est décroissante sur $\mathbb{R}_+$ et continue sur $\mathbb{R}_+^*$ d'après le théorème de continuité sous le signe $\int$.

2. Nous avons :

$$\begin{aligned} F(0) &= \int_0^{\frac{\pi}{2}} \frac{1 - \cos t}{\sin t} dt \\ &= \int_0^{\frac{\pi}{2}} \tan \frac{t}{2} dt = \ln 2. \end{aligned}$$

Par conséquent F est bornée. La continuité en 0 est une conséquence du théorème de convergence monotone appliqué à toute suite (x_n) décroissante et de limite 0.

De même en appliquant le théorème de convergence monotone à toute suite (x'_n) croissante de limite $+\infty$. En effet, pour tout t fixé, $f(x'_n, t)$ tend vers 0 sauf pour $t = \frac{\pi}{2}$ et la fonction limite est continue par morceaux sur $[0; \frac{\pi}{2}]$. Le théorème de convergence monotone sur les compacts entraîne que $F(x'_n)$ tend vers l'intégrale de la fonction caractéristique de $\frac{\pi}{2}$, soit 0. Donc $\lim_{x \rightarrow +\infty} F(x) = 0$.

3. Pour $x > 0$, la dérivation est possible selon le théorème de dérivation sous le signe $\int$ de Leibniz puisque la dérivée partielle est continue. Nous obtenons :

$$F'(x) = -\frac{1}{2} \int_0^{\frac{\pi}{2}} \frac{(1 - \cos t) \cos^2 t dt}{(\sin^2 t + x \cos^2 t)^{\frac{3}{2}}}.$$

Fiche 55

55.1

$$1. y(x) = \frac{1}{\cos(x)} \left(C + x - \tan\left(\frac{x}{2}\right) \right), C \in \mathbb{R}.$$

$$2. y(x) = C e^{2\epsilon(\arctan(\sqrt{\frac{1-x}{x}}) - \sqrt{x(1-x)})}, C \in \mathbb{R}, \epsilon = \pm 1.$$

$$3. y(x) = (C e^{\sqrt{x}} + x)^2, C \in \mathbb{R}.$$

Fiche 56

56.1

$$1. y(x) = \frac{x}{2} \sin(x) + \alpha \cos(x) + \beta \sin(x), \text{ avec } (\alpha, \beta) \in \mathbb{R}^2.$$

$$2. y(x) = (\alpha x + \beta + 16 \sin(x) - 8x \cos(x)) e^x + \frac{8}{125} ((15x + 4) \cos x - (20x + 22) \sin(x)) e^{-x}, \text{ avec } (\alpha, \beta) \in \mathbb{R}^2.$$

$$3. y(x) = e^{2x}(\alpha \cos(3x) + \beta \sin(3x)) + \frac{2988}{1305} \cos(2x) + \frac{156}{145} \sin(2x), \text{ avec } (\alpha, \beta) \in \mathbb{R}^2.$$

$$4. y(x) = \left(\frac{15}{12} x^2 + \frac{255}{36} x + \alpha \right) e^{2x} + \beta e^{-4x}, \text{ avec } (\alpha, \beta) \in \mathbb{R}^2.$$

$$5. y(x) = \alpha e^x + \beta e^{-2x} - \frac{8}{10} \cos(2x) - \frac{12}{5} \sin(2x), \text{ avec } (\alpha, \beta) \in \mathbb{R}^2.$$

$$6. y(x) = \alpha e^x + \left(\frac{-x^2}{6} - \frac{x}{9} + \beta \right) e^{-2x}, \text{ avec } (\alpha, \beta) \in \mathbb{R}^2.$$

Fiche 57

$$57.1 \quad \frac{-C}{y} - \ln|y| = x + C', (C, C') \in \mathbb{R}^2.$$

$$57.2 \quad dx = \frac{dy}{2\epsilon y \sqrt{1 + Cy^2}}.$$

$$\text{Si } C = 0, y = a e^{2\epsilon x}.$$

$$\text{Si } C < 0, y = \frac{k'}{\text{ch}(2(x + \lambda))}, (\lambda, k') \in \mathbb{R}^2, k' \neq 0.$$

$$\text{Si } C > 0, y = \frac{k''}{\text{sh}(2(x + \lambda'))}, (\lambda', k'') \in \mathbb{R}^2, k'' \neq 0.$$

$$57.3 \quad C \frac{\sin x}{x} + C' \frac{\cos x}{x}, (C, C') \in \mathbb{R}^2.$$

$$57.4 \quad y(x) = k_1 e^{\frac{x^2}{2}} + k_2 e^{-\frac{x^2}{2}}, (k_1, k_2) \in \mathbb{R}^2.$$

$$57.5 \quad y(x) = C_1 (x^3 + 9x) + C_2 (x^2 + 1), (C_1, C_2) \in \mathbb{R}^2.$$

$$57.6 \quad y(x) = k_1 x^2 \ln|x| + x^3 + k_2 x^2, (k_1, k_2) \in \mathbb{R}^2.$$

57.8

1. f est normalement convergente. f est définie sur $\mathbb{R}$, continue sur $\mathbb{R}$ et impaire. Elle est également $\mathcal{C}^1$ sur $\mathbb{R}^*$ et nous avons pour tout $x \in \mathbb{R}^*$:

$$f'(x) = \sum_{n=1}^{+\infty} \frac{1}{n(1+n^2x^2)} > 0.$$

f est strictement croissante sur $\mathbb{R}$.

2. Pour $x > 0$, l'encadrement

$$\arctan x \sum_{n=1}^{+\infty} \frac{1}{n^2} \leq f(x) \leq \frac{\pi}{2} \sum_{n=1}^{+\infty} \frac{1}{n^2}$$

permet de montrer que :

$$\lim_{x \rightarrow +\infty} f(x) = \frac{\pi^3}{12}.$$

3. La fonction $y \mapsto \frac{\arctan(xy)}{y^2}$ est décroissante sur $\mathbb{R}_+^*$. Un encadrement de la série à l'aide d'intégrales permet d'obtenir l'équivalent en 0 :

$$f(x) \sim -x \ln|x|.$$

57.9

1. f est définie sur $\mathbb{R}^*$ et impaire. Soit $x \in \mathbb{R}_+^*$, la fonction $t \mapsto \frac{1}{\text{sh}(tx)}$ est décroissante sur $\mathbb{R}_+^*$. Une comparaison série-intégrale donne l'équivalent en 0^+ :

$$f(x) \sim -\frac{\ln x}{x}.$$

2. Une comparaison série-intégrale donne aussi l'équivalent en $+\infty$:

$$f(x) \sim 2e^{-x}.$$

57.10 Si $x \leq 0$ la série est divergente et si $x > 0$ le critère de Riemann permet de montrer que la série converge.

Soit $x > 0$ fixé. La fonction $t \mapsto e^{-x\sqrt{t}}$ est décroissante sur $\mathbb{R}^+$ et donc une comparaison série-intégrale suivie d'une intégration par parties donne l'équivalent en 0^+ suivant :

$$f(x) \sim \frac{2}{x^2}.$$

57.11

1. Application du théorème de Leibniz.

$$2. \forall x \in \mathbb{R}_+, f'(x) = \frac{1}{1+x^2}.$$

$$\forall x \in \mathbb{R}_+, f(x) = \frac{\pi}{2} - \arctan x.$$

57.12

1. L'ensemble de définition est $\mathbb{R}$.

2. La continuité sur $\mathbb{R}$ provient du théorème de continuité sous le signe $\int$.

3. La classe $\mathcal{C}^1$ sur $\mathbb{R}^*$ provient du théorème de Leibniz de dérivation sous le signe $\int$.

4. Pour tout $x \neq 0$, nous avons :

$$F'(x) = 4x \int_0^{+\infty} \frac{dt}{(1+t^2)(x^2+t^2)}.$$

Un calcul permet de montrer que, pour tout $x \neq 0$:

$$F'(x) = \frac{2\pi}{x+1}.$$

Pour déterminer $F(0)$, il faut calculer $F(0) = 4 \int_0^{+\infty} \frac{\ln t}{1+t^2} dt$. En effectuant le changement de variable $u = \frac{1}{t}$, nous obtenons $F(0) = 0$.

5. Pour tout $x \in \mathbb{R}$, $F(x) = 2\pi \ln(|x|+1)$.

57.13

1. Le théorème de Leibniz s'applique, F est $\mathcal{C}^1$ sur $\mathbb{R}$, et ainsi pour tout $x \in \mathbb{R}$:

$$F'(x) = -2x \int_0^1 e^{-x^2(1+t^2)} dt.$$

2. La fonction intégrée est continue et ainsi le théorème sur les primitives des fonctions continues permet d'affirmer que G est $\mathcal{C}^1$ sur $\mathbb{R}$ et pour tout $x \in \mathbb{R}$:

$$G'(x) = 2e^{-x^2} \int_0^x e^{-t^2} dt.$$

3. En posant $u = xt$, nous montrons que $F' + G' = 0$. Donc $F + G$ est constante sur $\mathbb{R}$. Puis $F(0) + G(0) = \frac{\pi}{4}$.

4. Soit $x_0 \in \mathbb{R}$ fixé, nous avons :

$$\begin{aligned} |F(x_0)| &= \int_0^1 \frac{e^{-x^2(1+t^2)}}{1+t^2} dt \\ &\leq e^{-x^2} \int_0^1 \frac{1}{1+t^2} dt = \frac{\pi}{4e^{x^2}}. \end{aligned}$$

Par conséquent, comme $\frac{\pi}{4e^{x^2}} \rightarrow 0$ quand $x \rightarrow +\infty$, $\lim_{x \rightarrow +\infty} F(x) = 0$.

5. Pour $x > 0$, l'intégrale $\int_0^x e^{-t^2} dt \geq 0$ et ainsi nous avons :

$$\int_0^x e^{-t^2} dt = \sqrt{G(x)} = \sqrt{\frac{\pi}{2} - F(x)}.$$

En faisant tendre x vers $+\infty$ nous obtenons :

$$\int_0^{+\infty} e^{-t^2} dt = \frac{\sqrt{\pi}}{2}.$$

57.14 Soit f une solution au problème posé. f est dérivable sur $\mathbb{R}$ et comme $f'(x) = e^x - f(-x)$, f est deux fois dérivable sur $\mathbb{R}$. Dérivons l'égalité recherchée, pour tout $x \in \mathbb{R}$:

$$f''(x) = f'(-x) + e^x = -f(x) + e^{-x} + e^x.$$

f est donc solution sur $\mathbb{R}$ de l'équation différentielle $f''(x) + f(x) = 2 \operatorname{ch} x$. Par conséquent, il existe $(a, b) \in \mathbb{R}^2$ tel que, pour tout $x \in \mathbb{R}$:

$$f(x) = \operatorname{ch} x + a \cos x + b \sin x.$$

L'égalité initiale n'est vérifiée que si $a + b = 0$. Les solutions au problème posé sont donc les fonctions f telles que, pour tout $x \in \mathbb{R}$:

$$f(x) = \operatorname{ch} x + a(\cos x - \sin x)$$

où $a \in \mathbb{R}$.

57.15 La fonction nulle est solution. Considérons une solution non identiquement nulle. En utilisant la relation recherchée, il est facile de montrer que $f(0) = 0$ et que f est impaire. Comme f est continue, l'application $y \mapsto$

$\frac{1}{x_0} \int_{x_0-y}^{x_0+y} f(t)dt$ est de classe $\mathcal{C}^1$ donc f l'est aussi et ainsi $y \mapsto \frac{1}{x_0} \int_{x_0-y}^{x_0+y} f(t)dt$ est de classe $\mathcal{C}^2$. En dérivant l'égalité à y ou x fixé l'égalité recherchée nous obtenons pour tout $(x, y) \in \mathbb{R}^2$, $f'(x)f(y) = f(x+y) - f(x-y)$ et $f(x)f'(y) = f(x+y) + f(x-y)$. En dérivant la première égalité à y fixé et la seconde à x fixé, nous avons pour tout $(x, y) \in \mathbb{R}^2$:

$$f''(x)f(y) = f'(x+y) - f'(x-y) = f(x)f''(y).$$

En particulier f est solution d'une équation différentielle du type $y'' - ky = 0$. f étant impaire, f est de l'une des trois formes suivantes :

$$x \mapsto a \sin(bx), a \in \mathbb{R}^*, b > 0$$

$$x \mapsto ax, a \in \mathbb{R}^*$$

$$x \mapsto a \operatorname{sh}(bx), a \in \mathbb{R}^*, b > 0.$$

Réciproquement, et après avoir injecté les trois formes précédentes dans l'égalité recherchée, seules les fonctions suivantes conviennent :

$$x \mapsto \frac{2}{b} \sin(bx), b > 0$$

$$x \mapsto 2x$$

$$x \mapsto \frac{2}{b} \operatorname{sh}(bx), b > 0.$$

57.16

1. Les deux fonctions sont solutions de l'équation différentielle :

$$y'' + y = \frac{1}{x}.$$

Les deux fonctions ont la même limite lorsque $x \rightarrow +\infty$.

2. La continuité des deux fonctions en 0 donne :

$$\int_0^{+\infty} \frac{\sin t}{t} dt = \int_0^{+\infty} \frac{1}{1+t^2} dt = \frac{\pi}{2}.$$

57.18

1.a. Nous avons pour tout $t \in \mathbb{R}$, la majoration

$$\left| \frac{1}{(n!)^2} e^{int} \right| \leq \frac{1}{(n!)^2}.$$

Le membre de droite est le terme général d'une série convergente, ce qui nous assure de la

convergence normale, donc uniforme sur $\mathbb{R}$, de la série trigonométrique.

b. Le terme général, $t \mapsto \frac{1}{(n!)^2} e^{int}$, de la série trigonométrique est de classe $\mathcal{C}^2$ sur $\mathbb{R}$. Une majoration similaire à celle obtenue précédemment permet de montrer la convergence normale des séries des dérivées premières et secondes de $t \mapsto \frac{1}{(n!)^2} e^{int}$. Les théorèmes généraux permettent donc d'affirmer que f est de classe $\mathcal{C}^2$ sur $\mathbb{R}$ puis de vérifier que f est bien solution de l'équation différentielle (E).

2.a. Une double intégration par parties donne :

$$c_n(g'') = (in)^2 c_n(g) = -n^2 c_n(g).$$

b. g est solution de (E) : $g''(t) = -g(t)e^{it}$ pour tout $t \in \mathbb{R}$. Nous en déduisons :

$$\begin{aligned} c_n(g'') &= \frac{1}{2\pi} \int_{-\pi}^{\pi} g(t) e^{it} e^{-int} dt \\ &= -c_{n-1}(g). \end{aligned}$$

c. Les deux résultats précédents impliquent que, pour tout $n \in \mathbb{Z}$:

$$c_{n-1}(g) = n^2 c_n(g).$$

Deux récurrences simples donnent, pour tout

$$n \geq 0, c_n(g) = \frac{c_0(g)}{(n!)^2} \text{ et, pour tout } n < 0, c_n(g) = 0.$$

g étant de classe $\mathcal{C}^2$ sur $\mathbb{R}$, le théorème de Dirichlet s'applique et g est la somme de sa série de Fourier :

$$g(t) = \sum_{n \in \mathbb{Z}} c_n(g) e^{int} = c_0 \sum_{n \geq 0} \frac{1}{(n!)^2} e^{int}.$$

g est donc nécessairement proportionnelle à f : l'espace vectoriel des solutions 2π -périodiques est donc une droite vectorielle engendrée par f .

3. Le théorème de Cauchy indique que l'ensemble des solutions de (E) est un espace vectoriel de dimension 2. Il existe donc des solutions de (E) qui ne sont pas 2π -périodiques.

Fiche 58

58.1 Les lignes de niveau de f sont telles que $\frac{x^2}{4} + y^2 = c^2$, avec $c \in \mathbb{R}$. Si $c = 0$,

la « ligne » de niveau est l'origine $(0;0)$. Si $c \neq 0$, la ligne de niveau est une ellipse de grand axe $a = 2\sqrt{|c|}$ et de petit axe $b = \sqrt{|c|}$ car $\frac{x^2}{(2\sqrt{|c|})^2} + \frac{y^2}{\sqrt{|c|}^2} = 1$.

58.2 Les surfaces de niveau de f sont telles que $x^2 + y^2 + z^2 = c^2$ avec $c \in \mathbb{R}$. Si $c = 0$, la surface de niveau est l'origine $(0;0;0)$. Si $c \neq 0$, la surface de niveau est une sphère de rayon $|c|$.

58.3 Si f est continue sur le disque unité, elle est en particulier continue en l'origine $(0;0)$, donc $f(0;0) = \lim_{(x,y) \rightarrow (0;0)} f(x,y) = 7$.

58.4 Physiquement, la température ne peut varier brutalement d'une seconde à l'autre, ni d'une longitude y à une longitude très proche $y + \epsilon$, ni d'une latitude x à une latitude très proche $x + \epsilon$, donc la température est une fonction continue sur D .

58.5

1. La fonction f est continue sur $\mathbb{R}^2$, donc $\lim_{(x,y) \rightarrow (2;0)} f(x,y) = f(2;0) = 15$.

2. Prenons par exemple $x = 1/y$ avec $x \rightarrow 1$ et $y \rightarrow 1$. Alors $f(x, 1/x)$ n'est pas défini car le dénominateur est nul. La limite n'existe pas.

Fiche 59

59.1

1. Les dérivées partielles de r existent et sont continues sur $\mathbb{R}^3 \setminus \{\vec{0}\}$, r est donc différentiable sur $\mathbb{R}^3 \setminus \{\vec{0}\}$. Sa fonction gradient est $\nabla r(x,y,z) = \left(\frac{x}{\sqrt{x^2 + y^2 + z^2}}, \frac{y}{\sqrt{x^2 + y^2 + z^2}}, \frac{z}{\sqrt{x^2 + y^2 + z^2}} \right)$. La norme de la fonction gradient est égale à 1, et $\vec{\nabla} r = \vec{r}/\|\vec{r}\|$.

2. Pour $(x,y,z) \neq (0,0,0)$, $r^n = (x^2 + y^2 + z^2)^{n/2}$, $\frac{\partial r^n}{\partial x} = nx(x^2 + y^2 + z^2)^{n/2-1}$; $\frac{\partial r^n}{\partial y} = ny(x^2 + y^2 + z^2)^{n/2-1}$; $\frac{\partial r^n}{\partial z} = nz(x^2 + y^2 + z^2)^{n/2-1}$. C'est-à-dire $\vec{\nabla}(r^n) = nr^{n-2}\vec{r}$.

3. Ce sont les mêmes calculs qu'en 2.

4. Nous savons que $\vec{\nabla}(r^2) = 2\vec{r}$ d'après la question 2. Donc $f = r^2/2$ convient.

59.2

Si $(x,y) \neq (0,0)$, $\frac{\partial f}{\partial x}(x,y) = y^3/(x^2 + y^2)^{3/2}$ et $\frac{\partial f}{\partial y}(x,y) = x^3/(x^2 + y^2)^{3/2}$. Ces dérivées partielles sont continues sur $\mathbb{R}^2 \setminus \{(0,0)\}$, donc f est de classe $\mathcal{C}^1$ sur $\mathbb{R}^2 \setminus \{(0,0)\}$. En $(0,0)$, nous avons $\frac{\partial f}{\partial x}(0,0) = \lim_{x \rightarrow 0} \frac{f(x,0) - f(0,0)}{x} = 0$. D'autre part, $\lim_{x \rightarrow 0} \frac{\partial f}{\partial x}(x,x) = 1/2^{3/2} \neq 0$. Donc $\frac{\partial f}{\partial x}$ n'est pas continue en $(0,0)$, donc f n'est pas de classe $\mathcal{C}^1$ en $(0,0)$.

Fiche 60

60.1 La fonction $\vec{f}$ est bijective, elle est de classe $\mathcal{C}^1$ sur $]0; +\infty[\times]0; 2\pi[$ car ses dérivées partielles existent et sont continues sur $]0; +\infty[\times]0; 2\pi[$. De plus, le jacobien de $\vec{f}$ est égal à r , et ne s'annule donc jamais sur $]0; +\infty[\times]0; 2\pi[$. $\vec{f}$ est donc un $\mathcal{C}^1$ -difféomorphisme sur $]0; +\infty[\times]0; 2\pi[$.

60.2 La fonction $\vec{f}$ est de classe $\mathcal{C}^1$ sur $\mathbb{R}^3$ car ses dérivées partielles existent et sont continues sur $\mathbb{R}^3$. Cherchons (x,y,z) tel que $\vec{f}(x,y,z) = (x,y,z)$. Le seul point fixe est $(0;0;0)$.

60.3

1. Il faut montrer que $\frac{d\varphi(t, u(t))}{dt} = 0$, pour tout $t \in \mathbb{R}$.

2. Prenons $f(x,y) = (\sin x, y^2)$. Nous avons $f = \vec{\nabla} \varphi$ avec $\varphi(x,y) = -\cos x + \frac{y^3}{3}$. Nous devons avoir $\varphi(t, u(t)) = \text{constante}$, donc la solution u vérifie $-\cos(t) + u^3(t)/3 = \text{constante}$. D'où $u(t) = (\text{constante} + 3 \cos t)^{1/3}$. Comme $u(0) = 3$, nous en déduisons $u(t) = (24 + 3 \cos t)^{1/3}$.

60.4

1. Les dérivées partielles de f existent et sont continues sur $\mathbb{R}^2$ donc $f \in \mathcal{C}^1(\mathbb{R}^2)$.

4. Il faut utiliser le théorème des accroissements finis.

5. C'est le théorème du point fixe.

Fiche 61

61.1 $\frac{\partial^2 f}{\partial x^2} = 6x \cos(xy) - 6x^2 y \sin(xy) - x^3 y^2 \cos(xy), \frac{\partial^2 f}{\partial y^2} = -x^5 \cos(xy) + 6, \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} = -4x^3 \sin(xy) - x^4 y \cos(xy).$

$\vec{\nabla} f(0; 0) = (0; 0)$, donc $(0; 0)$ est un point stationnaire de f . Nous avons $f(0; 0) = 0$, et pour tout $\epsilon > 0$, $f(0; \epsilon) = 3\epsilon^2 > 0$ et $f(-\epsilon; 0) = -\epsilon^3 < 0$. Donc nous pouvons trouver des points arbitrairement proches de $(0; 0)$ tels que tantôt $f(0; 0)$ est dépassé, tantôt, $f(0; 0)$ n'est pas dépassé. $(0; 0)$ est donc un point selle.

61.2 f est de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$ et $\frac{\partial f}{\partial x} = 2(x+1)$ et $\frac{\partial f}{\partial y} = 2(y-1)$. f possède un unique point stationnaire qui est $(-1; 1)$. Nous avons $s^2 - rt = -4 < 0$ et $r = 2 > 0$ donc le point stationnaire est un minimum local. C'est même un minimum global car $f(x, y) \geq f(-1; 1) = 1$.

61.3 f est de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$ et $\frac{\partial f}{\partial x} = 2x$ et $\frac{\partial f}{\partial y} = -2y$. f possède un unique point stationnaire qui est $(0; 0)$. Nous avons $s^2 - rt = 4 > 0$ donc le point stationnaire est un point selle.

61.4 La surface du champ est xy , mais $2(x+y) = L$, donc $y = \frac{L}{2} - x$, et la surface à maximiser est $S(x) = x \times (\frac{L}{2} - x)$. S est de classe $\mathcal{C}^1$ et $S'(x) = \frac{L}{2} - 2x$. Le point stationnaire de S est $\frac{L}{4}$. Le tableau de variation de la fonction S montrer que ce point stationnaire est un maximum global sur $\mathbb{R}^+$. La surface maximale est donc atteinte pour une parcelle carrée de côté $\frac{L}{4}$, et la surface correspondante est de $\frac{L^2}{16} \text{ m}^2$.

61.5 f est continue sur C compact de $\mathbb{R}^2$, donc f est bornée et atteint ses bornes. Nous avons $\forall (x, y) \in C, -1 \leq f(x, y) \leq 1$, $f(1; 0) = 1$ et $f(0; 1) = -1$. Donc -1 et 1 sont respectivement les bornes inférieure et supérieure de f sur C .

Fiche 62

62.1 Soit $\vec{f}(t) = (x(t), y(t))$. $\vec{f}$ est de classe $\mathcal{C}^1$ sur $\mathbb{R}$ et $\|\vec{f}'(t)\| = 2R|\sin(t/2)|$, donc $L = \int_0^{2\pi} 2R \sin(t/2) dt = 8R$ est la longueur d'une arche de cycloïde.

Le rayon de courbure est $4R|\sin(t/2)| = 4R \sin(t/2)$ si $t \in [0; 2\pi]$.

62.2 Le rayon de courbure est $R(\theta) = \frac{3}{4}a \cos^2(\theta/3)$.

62.3 Nous avons $\|f'(t)\|^2 = \text{th}^2(t)$. La longueur de la courbe est $\int_0^t \text{th}(s) ds = [\ln(\text{ch } s)]_0^t = \ln(\text{ch } t)$.

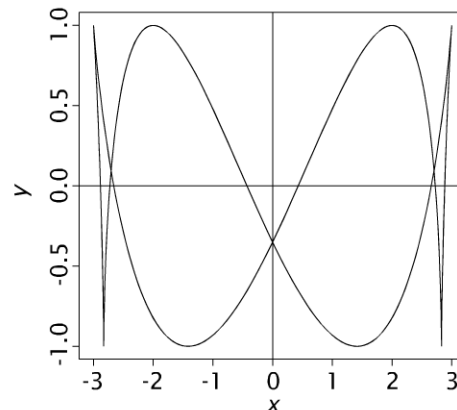
62.4

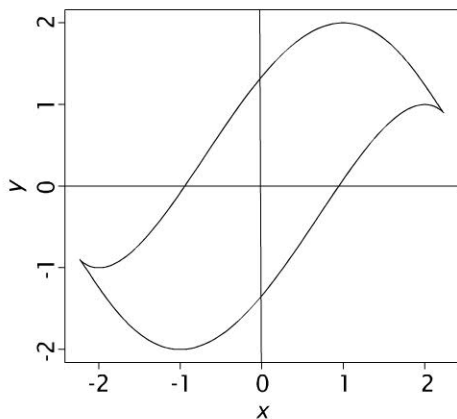
1. Prenons $\vec{f}(u, v) = x(u, v)\vec{i} + y(u, v)\vec{j} + z(u, v)\vec{k}$. $\vec{f}$ est de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$ et $\frac{\partial \vec{f}}{\partial u} \wedge \frac{\partial \vec{f}}{\partial v} = (R \cos u, R \sin u, 0)$, de norme égale à R en tout point (u, v) , donc l'aire du cylindre est $\int_0^{2\pi} \int_0^L R dv du = 2\pi RL$.

2. $\frac{\partial \vec{f}}{\partial u} \wedge \frac{\partial \vec{f}}{\partial v} = (2v \cos u, 2v \sin u, -v)$, de norme égale à $\sqrt{5}|v|$ en tout point (u, v) , donc l'aire du cône est $\int_0^{2\pi} \int_0^L \sqrt{5}v dv du = \sqrt{5}\pi L^2$.

Fiche 63

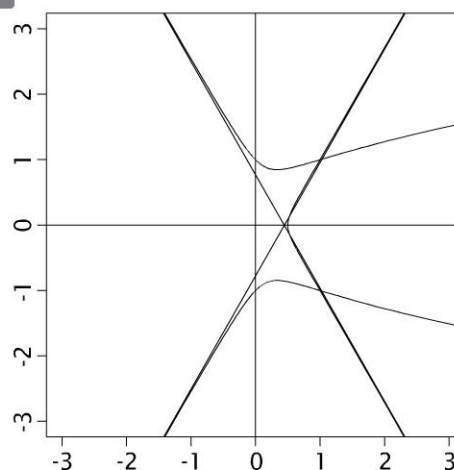
63.1





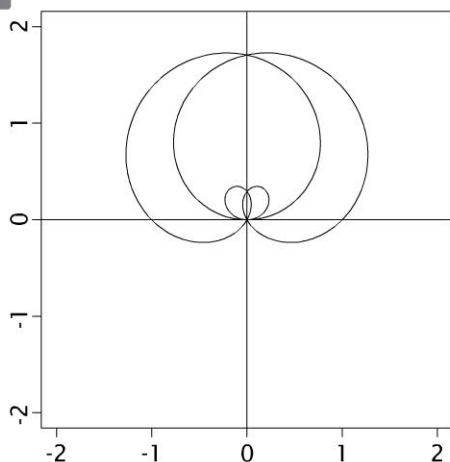
63.2 $x(t) = r \cos t, y(t) = r \sin(t)$ avec $t \in [0; 2\pi[$; $x(t) = a \cos t, y(t) = b \sin(t)$ avec $t \in [0; 2\pi[$

64.3

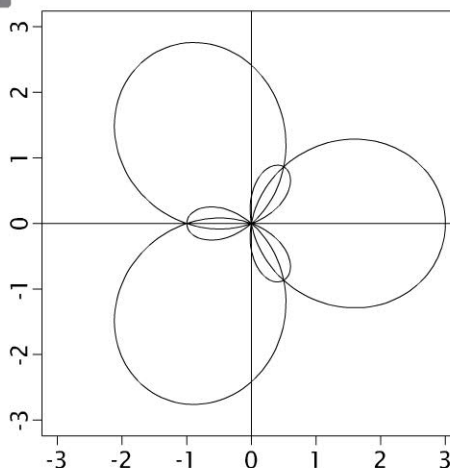


Fiche 64

64.1



64.2



Fiche 65

65.1

$$\begin{aligned} I &= \int_0^7 \left(\int_3^5 (x^2 - 7y^3) dy dx \right) \\ &= \int_0^7 \left[x^2 y - \frac{7}{4} y^4 \right]_3^5 dx = \int_0^7 (2x^2 - 952) dx \\ &= -\frac{19306}{3}. \end{aligned}$$

65.2

$$\begin{aligned} V &= \int \int_D (3x^2 + 2y^2) dx dy = \int_0^{\frac{3}{2}} \left(\int_{2x^2}^{3x} (3x^2 + 2y^2) dy \right) dx \\ &= \int_0^{\frac{3}{2}} \left[3x^2 y + \frac{2}{3} y^3 \right]_{2x^2}^{3x} dx. \end{aligned}$$

$$\text{Donc, } V = \int_0^{\frac{3}{2}} \left(27x^3 - 6x^4 - \frac{16x^6}{3} \right) dx = \frac{16767}{13440}.$$

65.3 $I = \frac{2}{3}.$

65.4 $I = \frac{51}{4} - \frac{21}{8} e^3.$

Fiche 66

66.1

1. Si $n \in \mathcal{E}_k$, alors $q_n = p_k$. Comme l'application $(n_1, \dots, n_k) \in \mathbb{N}^{k-1} \times \mathbb{N}^* \mapsto p_1^{n_1} \cdots p_k^{n_k} \in \mathcal{E}_k$ est une bijection, la famille $\left(\frac{1}{nq_n}\right)_{n \in \mathcal{E}_k}$ est sommable si et seulement si la famille $\left(\frac{1}{p_k p_1^{n_1} \cdots p_k^{n_k}}\right)_{(n_1, \dots, n_k) \in \mathbb{N}^{k-1} \times \mathbb{N}^*}$ est sommable ce qui est le cas puisque

$$\begin{aligned} & \frac{1}{p_k} \sum_{n_1=0}^{+\infty} \cdots \sum_{n_k=1}^{+\infty} \frac{1}{p_1^{n_1}} \cdots \frac{1}{p_k^{n_k}} \\ &= \frac{1}{p_k} \left(\sum_{n_1=0}^{+\infty} \frac{1}{p_1^{n_1}} \right) \cdots \left(\sum_{n_k=1}^{+\infty} \frac{1}{p_k^{n_k}} \right) \\ &= \frac{1}{p_k} \prod_{i=1}^{k-1} \left(\frac{1}{1 - \frac{1}{p_i}} \right) \frac{1}{p_k} \frac{1}{1 - \frac{1}{p_k}} = \frac{M_k}{p_k^2}. \end{aligned}$$

2. Posons $u_k = \frac{M_k}{p_k^2}$. Nous avons $\frac{p_{k+1}^3 u_{k+1}^2}{p_k^3 u_k^2} = \frac{p_{k+1}^3 p_k^4 p_{k+1}^2}{p_k^3 (p_{k+1} - 1)^2 p_{k+1}^4} = \frac{p_k p_{k+1}}{(p_{k+1} - 1)^2}$. Si $k \geq 2$, alors $p_k \leq p_{k+1} - 2$ et $p_k p_{k+1} \leq (p_{k+1} - 2)p_{k+1} < p_{k+1}^2 - 2p_{k+1} + 1 = (p_{k+1} - 1)^2$. La suite $(p_k^{3/2} u_k)_{k \geq 1}$ est donc bornée. Comme $p_k \geq k$, nous en déduisons que $k^{4/3} u_k$ tend vers 0 lorsque k tend vers $+\infty$. Le critère de Riemann permet d'affirmer que la série $\sum_{k=2}^{+\infty} u_k$ converge. Ainsi la série $\sum_{n=2}^{+\infty} \frac{1}{nq_n}$ converge et

$$\sum_{n=2}^{+\infty} \frac{1}{nq_n} = \sum_{k=1}^{+\infty} \left(\sum_{n \in \mathcal{E}_k} \frac{1}{nq_n} \right) = \sum_{k=1}^{+\infty} \frac{M_k}{p_k^2}.$$

66.2 Considérer la suite de compacts $K_n = \{(x, y) \mid 0 \leq x \leq 1, \frac{1}{n} \leq y^2 \leq x\}$ puis appliquer la formule de Fubini sur ces compacts. L'intégrale est égale à $\frac{\pi}{2}$.

66.3 La fonction f définie sur $[0; +\infty[$ par $f(x, y) = 4xye^{-x^2}e^{-y^2}$ est continue et bornée sur

$[0; +\infty[$. Le théorème de Fubini s'applique :

$$\begin{aligned} I &= \int_0^{+\infty} \left(\int_0^{+\infty} 4xye^{-x^2}e^{-y^2} dy \right) dx \\ &= \int_0^{+\infty} -2xe^{-x^2} \left(\int_0^{+\infty} -2ye^{-y^2} dy \right) dx \\ &= \int_0^{+\infty} -2xe^{-x^2} [e^{-y^2}]_0^{+\infty} dx \\ &= \int_0^{+\infty} -2xe^{-x^2} dx = [e^{-x^2}]_0^{+\infty} = 1. \end{aligned}$$

66.4

1. Introduire les disques $D_n = \{(x, y) \mid x^2 + y^2 \leq n^2\}$ puis passer en coordonnées polaires et utiliser la formule de Fubini :

$$\begin{aligned} u_n &= \iint_{D_n} \frac{dxdy}{(1 + x^2 + y^2)^\alpha} \\ &= \int_0^{2\pi} \left(\int_0^n \frac{rdr}{(1 + r^2)^\alpha} \right) d\theta \\ &= \pi \int_0^{n^2} \frac{dt}{(1 + t)^\alpha}. \end{aligned}$$

Ainsi l'intégrale double converge si et seulement si $\alpha > 1$.

2. $\left| \frac{\sin(x + y)}{(1 + x^2 + y^2)^2} \right| \leq \frac{1}{(1 + x^2 + y^2)^2}$ pour tout $(x, y) \in \mathbb{R}^2$. La fonction du second membre est intégrable sur $\mathbb{R}^2$ car il s'agit d'une fonction étudiée à la première question pour $\alpha = 2$. Le théorème de comparaison permet d'affirmer que l'intégrale étudiée est absolument convergente.

66.5

1. Le théorème de Fubini sur K_n permet de montrer que pour tout $n \in \mathbb{N}$, $u_n = \iint_{K_n} e^{-x^2-y^2} dxdy = \frac{\pi}{4} (1 - e^{-n^2})$. Par passage à la limite, nous obtenons : $K = \frac{\pi}{4}$.

2. Le théorème de Fubini sur L_n permet de montrer que pour tout $n \in \mathbb{N}$, $u_n = \iint_{L_n} e^{-x^2-y^2} dxdy = \int_0^n \left(\int_0^n e^{-x^2} e^{-y^2} dx \right) dy = \left(\int_0^n e^{-x^2} dx \right)^2$. Par passage à la limite, nous obtenons : $K = \left(\int_0^{+\infty} e^{-x^2} dx \right)^2$. Ainsi, nous obtenons : $\int_0^{+\infty} e^{-x^2} dx = \frac{\pi}{2}$.

66.6 Étudions l'intégrale sur les pavés $P_n = [-n; n] \times [-n; n]$:

$$\begin{aligned} u_n &= \int \int_{[-n;n] \times [-n;n]} \sin(x^2 + y^2) dx dy \\ &= 2 \int_0^{n^2} \frac{\sin t}{\sqrt{t}} dt \int_0^{n^2} \frac{\cos t}{\sqrt{t}} dt \end{aligned}$$

Le théorème d'Abel permet d'affirmer que ces deux intégrales convergent lorsque $n \rightarrow +\infty$.

Considérons maintenant les disques $D_n = \{(x, y) \mid x^2 + y^2 \leq n^2\}$:

$$\begin{aligned} v_n &= \int \int_{D_n} \sin(x^2 + y^2) dx dy \\ &= 2 \int_0^{2\pi} \left(\int_0^n \sin r^2 r dr \right) d\theta \\ &= \pi(1 - \cos n^2) \end{aligned}$$

qui est une suite qui diverge.

Si la fonction était intégrable, u_n et v_n devraient avoir la même limite, ce qui n'est pas le cas.

L'intégrale généralisée $\int \int_{\mathbb{R}^2} \sin(x^2 + y^2) dx dy$ est divergente.

Fiche 67

67.1 La courbe C est dans le plan, lisse par morceaux. Nous notons les fonctions $P(x, y) = x^5$ et $Q(x, y) = x^2 y$. Les fonctions P et Q sont de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$. D'après le théorème de Green dans le plan, nous avons : $\int_C (x^5 dx + x^2 y dy) = \int \int_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy$, où $D = \{(x, y) \in \mathbb{R}^2, \text{ tel que } 0 \leq x \leq 2, 0 \leq y \leq 3 - \frac{3x}{2}\}$ est l'intérieur du triangle parcouru

par C . Donc $I = \int_0^2 \int_0^{3-\frac{3x}{2}} (2xy - 0) dy dx = \int_0^2 x \left(3 - \frac{3x}{2} \right)^2 dx = 3$.

67.2 $\vec{F}(\vec{r}(t)) = \cos(t) \vec{i} - \cos^2(t) \sin(t) \vec{j}$, $\vec{r}'(t) = -\sin(t) \vec{i} + \cos(t) \vec{j}$. Donc, $\int_C \vec{F} \cdot d\vec{r} = 0$.

67.3 $P(x, y) = 2x + y$ et $Q(x, y) = 3x - y$. Les fonctions P et Q sont de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$ (qui est connexe par arcs) et $\frac{\partial P(x, y)}{\partial y} = 1$, $\frac{\partial Q(x, y)}{\partial x} = 3$. Comme $\frac{\partial P(x, y)}{\partial y} \neq \frac{\partial Q(x, y)}{\partial x}$, $\vec{F}$ n'est pas conservatif.

67.4 $I = 2\pi + \frac{2}{3}$.

67.5 $I = \frac{5\sqrt{5}-1}{6} + 2$.

67.6 $I = \frac{14}{3}$.

67.7 Nous remarquons que $\vec{F} = \vec{\nabla} \varphi$ avec $\varphi(x, y) = \frac{x^2 - 1}{y}$. La courbe C est incluse dans U qui est connexe par arcs, donc par le théorème sur la circulation d'un champ de gradient, on a : $\int_C \vec{F} \cdot d\vec{r} = \varphi(\vec{b}) - \varphi(\vec{a}) = \frac{b_1^2 - 1}{b_2} - \frac{a_1^2 - 1}{a_2}$.

67.8 Cherchons φ tel que $\vec{\nabla} \varphi = (P, Q)$ avec $P = -y/x^2 + y^2$ et $Q = x/x^2 + y^2$. Remarquons que $P = \frac{-1/y}{1 + (x/y)^2}$. Donc $\varphi(x, y) = -\arctan(x/y) + k(y)$; et $\partial_y \varphi = -\frac{1}{1 + (x/y)^2} \frac{-x}{y^2} + k'(y) = \frac{x}{y^2 + x^2} + k'(y) = Q(x, y)$. On en déduit que k est une constante. Donc $\varphi(x, y) = -\arctan(x/y)$ convient et est bien défini (y n'est jamais nul). Si la courbe C est dans le demi-plan supérieur (connexe par arcs), on peut appliquer le théorème sur la circulation d'un champ de gradient, et on a : $\int_C \frac{xdy - ydx}{x^2 + y^2} = \int_C \vec{\nabla} \varphi \cdot d\vec{r} = \varphi(\vec{b}) - \varphi(\vec{a}) = \arctan(a_1/a_2) - \arctan(b_1/b_2)$.

Si maintenant la courbe C n'est plus dans un ouvert connexe dans lequel (P, Q) est un gradient, on ne peut plus appliquer ce théorème ! C'est ce qui se passe avec le cercle trigonométrique : nous ne pouvons pas trouver d'ouvert connexe par arcs contenant le cercle trigonométrique, et tel que (P, Q) soit un gradient dans cet ouvert. Il faut alors calculer l'intégrale directement. Prenons la paramétrisation $x(t) = \cos t$, $y(t) = \sin t$, avec $t \in [0, 2\pi]$. Alors $\int_C \frac{xdy - ydx}{x^2 + y^2} = \int_0^{2\pi} \frac{\cos^2 t + \sin^2 t}{\cos^2 t + \sin^2 t} dt = 2\pi$.

Fiche 68

68.1 S se compose d'une face supérieure parabolique S_1 et d'une face de base ellipsoïdale S_2 . Pour S_1 , nous prenons la paramétrisation suivante : $\vec{r}(x, y) = x\vec{i} + y\vec{j} + g(x, y)\vec{k}$ où $g(x, y) = z = 1 - 2x^2 - y^2$, pour $(x, y) \in D = S_2$. Avec $P(x, y) = x$, $Q(x, y) = y$, $R(x, y) = z = 1 - 2x^2 - y^2$, et nous avons $\iint_{S_1} \vec{F} \cdot d\vec{S}$.

$d\vec{S} = \iint_D (2x^2 + y^2 + 1) dx dy$. Nous faisons le changement de variables $x = \frac{1}{\sqrt{2}}\rho \cos \theta$ et $y = \rho \sin \theta$, avec $(\theta, \rho) \in]0; 2\pi[\times]0; 1[$. Nous avons alors : $\iint_D (2x^2 + y^2 + 1) dx dy = \int_0^{2\pi} \int_0^1 (\rho^2 + 1) \frac{r}{\sqrt{2}} d\rho d\theta = \frac{3\pi}{2\sqrt{2}}$.

Pour S_2 , nous prenons la paramétrisation suivante : $\vec{r}(x, y) = x\vec{i} + y\vec{j}$ où $(x, y) \in D = S_2$. Ici, $z = g(x, y) = 0$. Avec $P(x, y) = x$, $Q(x, y) = y$, $R(x, y) = 0$, nous avons $\iint_{S_2} \vec{F} \cdot d\vec{S} = \iint_D 0 dx dy = 0$. Donc, $\iint_S \vec{F} \cdot d\vec{S} = \frac{3\pi}{2\sqrt{2}}$.

68.2 À l'intersection de la sphère et du cylindre, nous avons $z^2 = 8 \Rightarrow z = 2\sqrt{2}$, car $z > 0$.

S s'appuie sur le contour C décrit par les équations $x^2 + y^2 = 1$, $z = 2\sqrt{2}$. Une paramétrisation de C est donc donnée par $\vec{r}(t) = \cos(t)\vec{i} + \sin(t)\vec{j} + 2\sqrt{2}\vec{k}$, $0 \leq t \leq 2\pi$. Donc, $\vec{r}'(t) = -\sin(t)\vec{i} + \cos(t)\vec{j}$ et $\vec{F}(\vec{r}(t)) = 2\sqrt{2}\cos(t)\vec{i} + 2\sqrt{2}\sin(t)\vec{j} + \cos(t)\sin(t)\vec{k}$.

Nous avons alors par la formule de Stokes $\iint_S \text{rot} \vec{F} \cdot d\vec{S} = \int_C \vec{F} \cdot d\vec{r} = \int_0^{2\pi} \vec{F}(\vec{r}(t)) \cdot \vec{r}'(t) dt = 0$.

68.3 $I = \frac{\pi}{2}$.

68.4 Avec la représentation paramétrique $\vec{r}(u, v) = \sin u \cos v \vec{i} + \sin u \sin v \vec{j} + \cos u \vec{k}$, $I = \frac{\pi^2}{2}$.

Fiche 69

69.1 Pour tout $x \in \mathbb{R}_+$, $\mathcal{L}^{-1}[X](x) = \frac{2}{7}e^{-5x} + \frac{5}{7}e^{2x}$.

69.2 Pour tout $x \in \mathbb{R}_+$, $\mathcal{L}^{-1}[X](x) = \frac{1}{2}x \sin(x)$.

69.3 Pour tout $x \in \mathbb{R}_+$, $\mathcal{L}^{-1}[X](x) = -\frac{1}{15} + \frac{11}{140}e^{-5x} - \frac{1}{84}\cos(\sqrt{3}x) + \frac{11}{84}\sqrt{3}\cos(\sqrt{3}x)$.

69.4 La dérivée est $F'(p) = \frac{-2}{p(p^2 + 1)}$, pour tout $p \in \mathbb{R}_+^*$. La transformée de Laplace inverse de F' est égale $\mathcal{L}^{-1}[F'](x) = -2 + 2\cos(x)$, pour tout $x \in \mathbb{R}_+$. Pour obtenir l'original de F , nous utilisons la propriété qui indique que l'original de F' est égal à $-xf(x)$, pour tout $x \in \mathbb{R}_+$. Par conséquent, l'original de F est défini $f(x) = \frac{2(1 - \cos(x))}{x}$, pour tout $x \in \mathbb{R}_+^*$ et $f(0) = 0$. Remarquons que f est continue, et même dérivable, sur $\mathbb{R}_+$ car en 0 son développement limité est $f(x) = x + o(x)$.

Fiche 70

70.1 Soit le domaine $D = \{(x, y) | x + y > 0\}$. Le changement de variable défini par $x = r \cos \theta$ et $y = r \sin \theta$ est alors un $\mathcal{C}^1$ difféomorphisme de $\mathbb{R}_+^* \times]-\pi/4; 3\pi/4[$ sur D . En posant $F(r, \theta) = f(r \cos \theta, r \sin \theta)$, puis en différenciant cette fonction composée, nous avons :

$$\begin{aligned} \frac{\partial f}{\partial x} &= \frac{\partial F}{\partial r} \frac{\partial r}{\partial x} + \frac{\partial F}{\partial \theta} \frac{\partial \theta}{\partial x} \\ &= \cos \theta \frac{\partial F}{\partial r} - \frac{\sin \theta}{r} \frac{\partial F}{\partial \theta} \\ \frac{\partial f}{\partial y} &= \frac{\partial F}{\partial r} \frac{\partial r}{\partial y} + \frac{\partial F}{\partial \theta} \frac{\partial \theta}{\partial y} \\ &= \sin \theta \frac{\partial F}{\partial r} + \frac{\cos \theta}{r} \frac{\partial F}{\partial \theta}. \end{aligned}$$

En remplaçant dans l'équation de départ, nous obtenons $r \frac{\partial F}{\partial r} = (r^3 \cos^3(\theta) + r^3 \sin^3(\theta))^{1/2}$ ou encore :

$$\frac{\partial F}{\partial r} = \sqrt{r} \sqrt{\cos^3(\theta) + \sin^3(\theta)}.$$

Nous en déduisons que $F(r, \theta) = \frac{2}{3} \sqrt{\cos^3(\theta) + \sin^3(\theta)} r^{3/2} + \phi(\theta)$, où ϕ est une fonction dérivable sur $]-\pi/4; 3\pi/4[$.

En utilisant les variables de départ, nous avons :

$$f(x, y) = \frac{2}{3} \sqrt{x^3 + y^3} + \psi(x, y)$$

où ψ est différentiable et la valeur de $\psi(x, y)$ ne dépend que de celle du rapport $\frac{y}{x}$.

70.3 Soit $\phi : (x, y) \mapsto (u, v)$ avec $u = x^2 - y^2 - 2xy$ et $v = y$. Posons $\Delta_1 = \{(x, y) \in \mathbb{R}^2 | x > y\}$, $\Delta_1 = \{(x, y) \in \mathbb{R}^2 | x < y\}$ et $D = \{(u, v) \in \mathbb{R}^2 | u + 2v^2 > 0\}$.

ϕ réalise un $\mathcal{C}^1$ difféomorphisme de D_1 , ou D_2 , sur D , la bijection réciproque étant définie par $(u, v) \in D \mapsto (v + \sqrt{u + 2v^2}, v) \in \Delta_1$, ou par $(u, v) \in D \mapsto (v - \sqrt{u + 2v^2}, v) \in \Delta_2$.

Sur Δ_1 ou sur Δ_2 , procédons au changement de variable dans l'équation aux dérivées partielles en posant $f(u, v) = F(u(x, y), v(x, y))$ d'où :

$$\begin{aligned}\frac{\partial f}{\partial x} &= \frac{\partial F}{\partial u} \frac{\partial u}{\partial x} + \frac{\partial F}{\partial v} \frac{\partial v}{\partial x} \\ &= 2(x - y) \frac{\partial F}{\partial u} \\ \frac{\partial f}{\partial y} &= \frac{\partial F}{\partial u} \frac{\partial u}{\partial y} + \frac{\partial F}{\partial v} \frac{\partial v}{\partial y} \\ &= -2(x + y) \frac{\partial F}{\partial u} + \frac{\partial F}{\partial v}.\end{aligned}$$

En remplaçant dans l'équation de départ, nous obtenons :

$$\begin{aligned}(x + y) \frac{\partial f}{\partial x} + (x - y) \frac{\partial f}{\partial y} &= 0 \\ \Downarrow \\ \frac{\partial F}{\partial v} &= 0.\end{aligned}$$

Nous en déduisons que $F(u, v) = \phi(u)$, où ϕ est une fonction dérivable sur $\mathbb{R}$.

En utilisant les variables de départ, nous avons :

$$f(x, y) = \psi(x^2 - y^2 - 2xy, y)$$

où ψ est différentiable et constante par rapport à la seconde variable y .

70.4 ϕ existe le changement de variable linéaire, donc licite, $t = \theta - \pi$ permet de montrer que ϕ est paire. Une dérivation sous le signe $\int$, puis une application de la formule de Green-Riemann donne que $r\phi'(r) = 0$ pour tout $t \in \mathbb{R}^+$. Par continuité, ϕ' est nulle sur $\mathbb{R}_+$ puis par parité ϕ' est nulle sur $\mathbb{R}$. Ainsi ϕ est constante sur $\mathbb{R}$ et en utilisant la valeur $r = 0$, nous obtenons $\phi(r) = \phi(0) = 2\pi f(0, 0)$, pour tout $r \in \mathbb{R}$.

Fiche 71

71.1

Soit N le nombre de plaques autorisées par ce système : $N = (23^2 - 2) \times 999 \times (23^2 - 1) = 277\,977\,744$.

71.2

1. Il y a $A_{22}^{10} = 2,35 \times 10^{12}$ listes ordonnées des dix premiers pilotes à l'arrivée.

2. Il y a $\binom{22}{10} = 646\,646$ listes non ordonnées des dix premiers pilotes à l'arrivée.

71.3

1. Il s'agit de donner des listes ordonnées de 3 élèves (car les trois fonctions au bureau sont différentes : le triplet d'élèves « Paul président, Julie trésorière, Dominique secrétaire » n'est pas le même que le triplet des trois mêmes élèves « Julie présidente, Paul trésorier, Dominique secrétaire »). Il y a $A_{500}^3 = 124\,251\,000$ bureaux possibles.

2. Il s'agit ici d'ordonner 20 élèves. Le nombre de listes ordonnées de ces 20 élèves est le nombre de permutations des 20 élèves : $20! \approx 2,43 \times 10^{18}$.

71.4 Il s'agit de donner des listes ordonnées de 4 demoiselles. Il y a $A_{30}^4 = 657\,720$ classements possibles.

71.5 Pour former un acide aminé avec trois bases azotées, nous piochons avec remise parmi les quatre bases possibles. Ceci nous donne $4^3 = 64$ acides aminés possibles.

Fiche 72

72.1

Si $(A_i)_{i=1,\dots,n}$ est une partition de Ω , alors $\cup_{i=1}^n A_i = \Omega$, donc $\mathbb{P}(\cup_{i=1}^n A_i) = \mathbb{P}(\Omega) = 1$.

72.2

Nous avons $A \cap B = \emptyset$, donc A et B sont disjoints, donc $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) = 0,7$.

72.3

Nous avons $A \cap B \subset A$ donc $\mathbb{P}(A \cap B) \leq \mathbb{P}(A)$. De plus, $\mathbb{P}(A \cap B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cup B) \geq 0,6 + 0,7 - 1 = 0,3$. D'où $0,3 \leq \mathbb{P}(A \cap B) \leq 0,6$.

72.4 $\mathbb{P}(\text{« voter blanc »}) = 0,4$.

72.5 $\Omega = \{1; \dots; 6\} \times \{1; \dots; 6\}$, et $\text{Card}(\Omega) = 6^2 = 36$.

72.6 $\Omega_1 = \{1; \dots; 10\} \times \{1; \dots; 10\}$, et $\text{Card}(\Omega_1) = 10^2 = 100$.

Ω_2 est l'ensemble des binômes de boules prélevés simultanément dans l'urne contenant 10 boules, $\text{Card}(\Omega_2) = \binom{10}{2} = 45$.

72.7 $A \subset B$ donc $\mathbb{P}(A) \leq \mathbb{P}(B)$.

72.8 $\mathbb{P}(C) = \mathbb{P}(\bar{C}) = 0,5$ et $\mathbb{P}(A_2 \cup A_3 \cup A_4) = 0,75$. $B \subset \bar{C}$ donc $\mathbb{P}(B) \leq \mathbb{P}(\bar{C})$.

72.9 $\mathbb{P}(\bar{A} \cap \bar{B}) - \mathbb{P}(\bar{A})\mathbb{P}(\bar{B}) = \mathbb{P}(\bar{A}) + \mathbb{P}(\bar{B}) - \mathbb{P}(A \cup B) - (1 - \mathbb{P}(A)) \times (1 - \mathbb{P}(B)) = 2 - \mathbb{P}(A) - \mathbb{P}(B) - (1 - \mathbb{P}(A \cap B)) - (1 - \mathbb{P}(A)) \times (1 - \mathbb{P}(B)) = \mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)$

Fiche 73

73.1 $p = \frac{1}{12}$. La probabilité d'un tirage pair est $\frac{7}{12}$.

73.2 Ici, Ω est l'ensemble de toutes les listes possibles non ordonnées de six nombres tirés sans remise parmi $\{1; \dots; 49\}$. Ω est muni de la probabilité uniforme. La probabilité de tirer les six bons numéros est égale à $1/\text{Card}(\Omega) = 1/\binom{49}{6} \approx 7,15 \times 10^{-8}$.

73.3 Soit $\Omega = \{1; \dots; 365\}^n$ l'ensemble des listes possibles des dates d'anniversaire des n élèves. Ω est muni de la probabilité uniforme. Soit $A = \text{« au moins deux élèves de la classe sont nés le même jour »}$. Si $n > 365$, alors $\mathbb{P}(A) = 1$. Si $n \leq 365$, alors $\mathbb{P}(A) = 1 - \mathbb{P}(\bar{A}) = 1 - \frac{\text{Nombre de cas où les élèves sont nés à des dates distinctes}}{\text{Card}(\Omega)} = 1 - \frac{A_n}{365^n}$.

73.4 Il faut $1 = \sum_{k=1}^{+\infty} p_k = \alpha(e^1 - 1)$, donc $\alpha = 1/(e^1 - 1)$. Nous avons utilisé les résultats sur la série exponentielle.

73.5 Il faut $1 = \int_{\mathbb{R}} f(x) dx = C \int_0^{+\infty} x e^{-x^2} dx = C/2$, d'où $C = 2$. La pro-

babilité de l'intervalle $[-2; 2]$ est $\int_{-2}^2 f(x) dx = \int_0^2 2x e^{-x^2} dx = 1 - e^{-4}$. De même, la probabilité de l'intervalle $[2; +\infty[$ est égale à $\int_2^{+\infty} f(x) dx = e^{-4}$, les probabilités des intervalles $] - 2; 2]$, $[-2; 2[$ et $] - 2; 2[$ sont les mêmes que celle de l'intervalle $[-2; 2]$, et la probabilité de l'intervalle $]2; +\infty]$ est la même que celle de l'intervalle $[2; +\infty[$.

73.6 Ici Ω est la cible de rayon 2 m, muni de la probabilité uniforme. Soit A l'événement « la cible tombe à moins de 1 m du centre de la cible ». Alors $\mathbb{P}(A) = \frac{\text{Aire de } A}{\text{Aire de } \Omega} = \frac{1}{4}$.

73.7 Ici Ω est le carré de côté 1, muni de la probabilité uniforme. Alors $\mathbb{P}(B) = \frac{\text{Aire de } B}{\text{Aire de } \Omega}$. Nous pouvons aisément constater sur le graphique 73.1 que l'aire de B est égale à $\frac{1}{2} - \frac{1}{2} \times \frac{\pi}{4} \approx 0,107$.

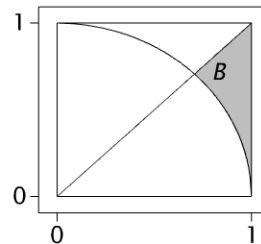


Figure 73.1 – Calcul de l'aire de B , Exercice 73.7.

Fiche 74

74.1 Soit l'événement $P = \text{« le système est en panne »}$. Nous supposons que les deux résistances sont indépendantes.

1. En parallèle, $\mathbb{P}(P) = \mathbb{P}(R_1 \cap R_2) = p_1 p_2$. En série, $\mathbb{P}(P) = \mathbb{P}(R_1 \cup R_2) = p_1 + p_2 - p_1 p_2$. Dans les deux cas, nous avons utilisé l'indépendance des événements R_1 et R_2 .

2. En parallèle, $\mathbb{P}(R_1|P) = 1$. En série, $\mathbb{P}(R_1|P) = \mathbb{P}(R_1 \cap P)/\mathbb{P}(P) = p_1/(p_1 + p_2 - p_1 p_2)$.

74.3 Soient les événements R = « réussir au Numérus Clausus », et D = « être un(e) redoublant(e) ». Nous utilisons la Formule des probabilités totales, pour calculer le taux de réussite au Numérus Clausus : $\mathbb{P}(R) = \mathbb{P}(R|D)\mathbb{P}(D) + \mathbb{P}(R|\bar{D})\mathbb{P}(\bar{D}) = 0,8 \times \frac{3}{10} + 0,3 \times \frac{7}{10} = 0,45$.

74.4 Soient les événements e_i = « être étudiant(e) de l'école i », pour $i = 1; \dots; 3$, et F = « être une étudiante du campus ». Le campus compte 450 étudiants. Nous avons $\mathbb{P}(e_1) = \frac{140}{450}$, $\mathbb{P}(e_2) = \frac{160}{450}$ et $\mathbb{P}(e_3) = \frac{150}{450}$. D'après la Formule des probabilités totales, le pourcentage d'étudiantes dans le campus est $\mathbb{P}(F) = \sum_{i=1}^3 \mathbb{P}(F|e_i) \times \mathbb{P}(e_i) = 0,1 \times 140/450 + 0,8 \times 160/450 + 0,5 \times 150/450 \approx 0,48$. Nous calculons ensuite $\mathbb{P}(e_1|F)$ par le Théorème de Bayes : $\mathbb{P}(e_1|F) = \frac{\mathbb{P}(F|e_1)\mathbb{P}(e_1)}{\mathbb{P}(F)} = \frac{0,1 \times 140/450}{0,1 \times 140/450 + 0,8 \times 160/450 + 0,5 \times 150/450} \approx 0,064$.

Fiche 75

75.1 $G(\Omega) = \{-1; +1\}$, $\mathbb{P}(G = -1) = \frac{19}{37}$ et $\mathbb{P}(G = +1) = \frac{18}{37}$.

75.2 Supposons que les trois dés sont non truqués. Les résultats des tirages sont donc équiprobables. $S(\Omega) = \{3; 4; 5; \dots; 18\}$. L'événement $\{S = 7\}$ est réalisé si les faces $\{1; 1; 5\}$, $\{1; 3; 3\}$, $\{2; 2; 3\}$ ou $\{1; 2; 4\}$ sortent au tirage des trois dés. Chacun des trois premiers cas cités peut être atteint de trois manières différentes, par exemple $(1; 1; 5)$, ou $(1; 5; 1)$ ou $(5; 1; 1)$. Le dernier cas peut être obtenu de $3! = 6$ manières différentes, correspondant au nombre de permutations des chiffres $\{1; 2; 4\}$. L'événement $\{S = 7\}$ peut donc survenir de 15 manières différentes, et nous avons un total de 6^3 résultats possibles. Donc, $\mathbb{P}(S = 7) = \frac{15}{6^3}$. De même, nous trouvons $\mathbb{P}(S = 8) = \frac{21}{6^3}$.

75.4

- Nous avons $1 = \int_{\mathbb{R}} f_X(x) dx$, d'où $A = \frac{1}{\pi}$.
- La probabilité est $\mathbb{P}(X \in [-1; 1]) = \int_{-1}^1 f_X(x) dx = \frac{1}{2}$.

75.5

- Nous avons $1 = \int_{\mathbb{R}} f_X(x) dx$, d'où $A = 2h^2$.
- La probabilité que la fusée se pose à moins de 10 m de la cible est $\mathbb{P}(X < 10) = \int_0^{10} f_X(x) dx = 1 - e^{-100h^2} = 0,90$, nous en déduisons $h \approx 0,15$.

Fiche 76

76.1

$F_X(x) = \frac{1}{6} \mathbf{1}_{x \in [1; 2[} + \frac{1}{3} \mathbf{1}_{x \in [2; 3[} + \frac{1}{2} \mathbf{1}_{x \in [3; 4[} + \frac{2}{3} \mathbf{1}_{x \in [4; 5[} + \frac{5}{6} \mathbf{1}_{x \in [5; 6[} + \mathbf{1}_{x \geq 6}$. C'est une fonction en escaliers.

76.2

- $X(\Omega) = \{3; 5; 8\}$; $\mathbb{P}(X = 3) = 0,3$; $\mathbb{P}(X = 5) = 0,3$ et $\mathbb{P}(X = 8) = 0,4$.
- , 3., 4., 5., 6. et 7. Ensuite, $\mathbb{P}(X < 2) = 0$; $\mathbb{P}(X > 3/2) = 1$; $\mathbb{P}(3 < X < 5) = 0$; $\mathbb{P}(3 \leq X < 5) = 0,3$; $\mathbb{P}(3 < X \leq 5) = 0,3$; $\mathbb{P}(3 \leq X \leq 5) = 0,6$.

76.3

- $F_X(x) = \sum_{k \leq x} p_k$.
- La probabilité que trois forages au moins donnent du pétrole est $\mathbb{P}(X \geq 3) = 1 - \mathbb{P}(X < 3) = 1 - \mathbb{P}(X \leq 2) = 1 - p_0 - p_1 - p_2$.

76.4

- Nous avons $0,90 = \mathbb{P}(X > 10) = 1 - F_X(10) = e^{-10\lambda}$, d'où $\lambda = \frac{-\ln(0,90)}{10} \approx 0,0105$.
- $\mathbb{P}(X > 5) = 1 - F_X(5) = e^{-5\lambda} \approx 0,9487$.
- $\mathbb{P}(X > 15|X > 10) = \frac{\mathbb{P}(X > 10 \text{ et } X > 15)}{\mathbb{P}(X > 10)} = \frac{e^{-15\lambda}}{e^{-10\lambda}} = e^{-5\lambda} = \mathbb{P}(X > 5)$: le fait que la batterie a déjà vécu 10 unités de temps n'entre pas en

jeu dans le résultat. La batterie est dite « sans vieillissement ».

5. La densité de X est donnée par $f_X(x) = F'_X(x) = \lambda e^{-\lambda x} \mathbb{1}_{x \geq 0}$.

76.5

$F_X(x) = \int_{-\infty}^x f_X(x) dx$ ce qui donne $F_X(x) = 0$ si $x \leq 0$ et $F_X(x) = 1 - e^{-h^2 x^2}$ si $x > 0$. Au final, la fonction de répartition est donc $F_X(x) = (1 - e^{-h^2 x^2}) \mathbb{1}_{x \geq 0}$.

Fiche 77

77.1

$X(\Omega) = \{a\}$ et $p_X(a) = 1$, donc $\mathbb{E}(X) = a \times p_X(a) = a$.

77.2

$$\mathbb{E}(X) = \frac{1}{2}f_1 + \frac{1}{8}f_2 + \frac{3}{8}f_3.$$

77.3

La variable aléatoire $\mathbb{1}_{X \in A}$ prend les valeurs 1 et 0, avec probabilités respectives $\mathbb{P}(X \in A)$ et $\mathbb{P}(X \notin A)$. Donc $\mathbb{E}(X) = 1 \times \mathbb{P}(X \in A) + 0 \times \mathbb{P}(X \notin A) = \mathbb{P}(X \in A)$.

77.4

Soit X le gain de l'organisme. Nous avons $X(\Omega) = \{2; 2 - c\}$, $\mathbb{P}(X = 2) = 1 - p$ et $\mathbb{P}(X = 2 - c) = p$, d'où $\mathbb{E}(X) = 2 - pc$. Il faut que $c < \frac{2}{p}$ pour que $\mathbb{E}(X) > 0$, c'est-à-dire une cagnotte inférieure à 4 si $p = \frac{1}{2}$, à 20 si $p = \frac{1}{10}$ et à 20 millions si $p = \frac{1}{10^7}$.

77.5

$$\mathbb{E}(X) = 0,05 \times \sum_{i=1}^3 \mathbb{E}(X_i) + 0,85 \times \mathbb{E}(X_4).$$

Fiche 78

78.1

Le taux de diabète doit varier peu, donc, il est préférable d'avoir la plus faible variance. La personne ayant une variance de 1 doit se faire soigner.

$$\begin{aligned} 78.2 \quad \mathbb{E}(X) &= 0,50 \times 0,1 + 0,20 \times 0,7 + 0,10 \times 0,2 = 0,21 \text{ et } \mathbb{V}\text{ar}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \\ &= 0,50^2 \times 0,1 + 0,20^2 \times 0,7 + 0,10^2 \times 0,2 - (0,21)^2 = 0,109. \end{aligned}$$

78.3

La propriété de linéarité de l'espérance donne $\mathbb{V}\text{ar}(X) = \mathbb{E}((X - \mathbb{E}(X))^2) = \mathbb{E}(X^2 - 2\mathbb{E}(X)X + \mathbb{E}(X)^2) = \mathbb{E}(X^2) - 2\mathbb{E}(X)^2 + \mathbb{E}(X)^2 = \mathbb{E}(X^2) - \mathbb{E}(X)^2$. La variance est positive car c'est l'espérance de $(X - \mathbb{E}(X))^2$, qui est une variable aléatoire positive.

78.4

Nous avons $X = \mathbb{E}(X) = a$, donc $\mathbb{V}\text{ar}(X) = \mathbb{E}((X - \mathbb{E}(X))^2) = \mathbb{E}((a - a)^2) = 0$.

78.5

$\mathbb{V}\text{ar}(aX + b) = \mathbb{E}((aX + b - \mathbb{E}(aX + b))^2) = \mathbb{E}((aX - \mathbb{E}(aX))^2) = a^2 \mathbb{V}\text{ar}(X)$ par linéarité de l'espérance. Si $\mathbb{E}(X) = \mu$ et $\mathbb{V}\text{ar}(X) = \sigma^2$, alors $\mathbb{V}\text{ar}\left(\frac{X - \mu}{\sigma}\right) = \frac{1}{\sigma^2} \mathbb{V}\text{ar}(X) = 1$. Enfin, la propriété de linéarité de l'espérance donne $\mathbb{E}\left(\frac{X - \mu}{\sigma}\right) = \frac{1}{\sigma}(\mathbb{E}(X) - \mu) = 0$.

78.6

Soient f_X et f_Y les densités des variables aléatoires X et Y . Si X et Y ont même loi, alors $f_X = f_Y$. Elles ont donc la même espérance car $\mathbb{E}(X) = \int_{\mathbb{R}} x f_X(x) dx = \int_{\mathbb{R}} y f_Y(y) dy = \mathbb{E}(Y) = \mu$. Leurs moments s'expriment par $\tilde{\mu}_r = \mathbb{E}((X - \mathbb{E}(X))^r) = \int_{\mathbb{R}} (x - \mathbb{E}(X))^r f_X(x) dx = \int_{\mathbb{R}} (x - \mu)^r f_X(x) dx = \int_{\mathbb{R}} (y - \mu)^r f_Y(y) dy$: les moments sont identiques.

Fiche 79

79.1

X n'a pas de mode, car toutes les valeurs de $X(\Omega)$ sont équiprobables.

79.2

1. La durée de vie médiane est la valeur m telle que $F_X(m) = 0,5$, soit $m = \frac{-\ln(0,5)}{\lambda}$.
2. La valeur 10 est le quantile d'ordre 10 % car $\mathbb{P}(X \leq 10) = 0,10$ et $\lambda = -\ln(0,9)/10$.

3. Un intervalle de probabilité de niveau 95 % est donné par $[x_{0,025}; x_{0,975}]$ où x_p désigne le quantile d'ordre p de la variable aléatoire X . Nous obtenons x_p en résolvant $F_X(x_p) = p$, ce qui nous donne l'intervalle $\left[\frac{-\ln(0,975)}{\lambda}; \frac{-\ln(0,025)}{\lambda} \right]$.

79.3 1. 0 2. 0 3. -1,960

4. $[-1,960; 1,960]$

Fiche 80

80.1 $Y(\Omega) = \{a + 3b; a + 5b; a + 8b\}$ et à condition que les trois valeurs de $Y(\Omega)$ soient distinctes, nous avons $\mathbb{P}(Y = a + 3b) = 0,3$; $\mathbb{P}(Y = a + 5b) = 0,3$ et $\mathbb{P}(Y = a + 8b) = 0,4$.

80.2 Soit F_S la fonction de répartition de S . Si $s < 0$, $F_S(s) = 0$. Si $s \geq 0$, $F_S(s) = \mathbb{P}(\pi R^2 \leq s) = \mathbb{P}\left(|R| \leq \sqrt{\frac{s}{\pi}}\right) = \mathbb{P}\left(R \leq \sqrt{\frac{s}{\pi}}\right)$ car R est à valeurs positives. Nous obtenons $F_S(s) = F_R\left(\sqrt{\frac{s}{\pi}}\right) = 1 - e^{-h^2 s/\pi}$.

La densité f_S de S est alors $f_S(s) = F'_S(s) = \frac{h^2}{\pi} e^{-h^2 s/\pi} \mathbb{1}_{s \geq 0}$.

80.3 Soit $Y = \left\lfloor \frac{X}{3} \right\rfloor$ où $\lfloor \cdot \rfloor$ désigne la partie entière. Nous avons $Y(\Omega) = \mathbb{N}$, $\mathbb{P}(Y = 0) = \mathbb{P}(X < 3) = \int_0^3 f_X(x) dx$, $\mathbb{P}(Y = 1) = \mathbb{P}(3 < X \leq 6) = \int_3^6 f_X(x) dx$. De même, pour tout $k \in \mathbb{N}$, $\mathbb{P}(Y = k) = \int_{3k}^{3(k+1)} f_X(x) dx$.

80.4

1. $X(\Omega) = \{-1; x\}$, $\mathbb{P}(X = -1) = \frac{19}{37}$ et $\mathbb{P}(X = x) = \frac{18}{37}$. $\mathbb{E}(X) = \frac{1}{37}(18x - 19) > 0$ si $x > \frac{19}{18} \approx 1,055$.

2. $Y(\Omega) = \{-1; y\}$, $\mathbb{P}(Y = -1) = \frac{36}{37}$ et $\mathbb{P}(Y = y) = \frac{1}{37}$. $\mathbb{E}(Y) = \frac{1}{37}(y - 36) > 0$ si $y > 36$.

3. $Z(\Omega) = \{-3; -2 + z; -1 + 10z\}$, $\mathbb{P}(Z = -3) = \frac{18}{37}$, $\mathbb{P}(Z = -2 + z) = \frac{18}{37}$ et $\mathbb{P}(Z = -1 + 10z) = \frac{1}{37}$. $\mathbb{E}(Z) = \frac{1}{37}(28z - 91) > 0$ si $z > \frac{91}{28} \approx 3,25$.

80.5

1. $\forall x \in \mathbb{R}, F_X(x) = (1 - e^{-\lambda x}) \mathbb{1}_{x \geq 0}$;

2. $\forall u \in [0; 1[, G(u) = -\frac{\ln(1 - u)}{\lambda}$;

3. $G(U)$ et X ont même loi car $\forall x \in \mathbb{R}$, $\mathbb{P}(G(U) \leq x) = F_X(x)$.

80.6 La variable aléatoire G est à valeurs positives. Soit $g \geq 0$. $\mathbb{P}(G \leq g) = \mathbb{1}_{\beta y \leq g} \times (1 - F_X(y)) + F_X\left(\min\left(y; \frac{g + y\delta}{\beta + \delta}\right)\right)$. G est une variable aléatoire dite « mixte ».

80.7 La variable aléatoire R est à valeurs dans $R(\Omega) = \left[\frac{v^2}{g} \sin\left(\frac{\pi}{2} - \frac{\pi}{180}\right); \frac{v^2}{g} \right]$.

• Soit $r \in R(\Omega)$, alors $\mathbb{P}(R \leq r) = \mathbb{P}\left(\alpha \in \left[\frac{\pi}{4} - \frac{\pi}{360}; \alpha_0\right] \cup \left[\alpha_1; \frac{\pi}{4} + \frac{\pi}{360}\right]\right)$ où $\alpha_0 = \frac{1}{2} \text{Arcsin}\left(\frac{gr}{v^2}\right)$ et $\alpha_1 = \frac{\pi}{2} - \alpha_0$. Donc, $\mathbb{P}(R \leq r) = \frac{180}{\pi} \left(\text{Arcsin}\left(\frac{gr}{v^2}\right) - \frac{\pi}{2} + \frac{\pi}{180} \right) = \frac{180}{\pi} \text{Arcsin}\left(\frac{gr}{v^2}\right) - 90 + 1$.

• Si $r < \frac{v^2}{g} \sin\left(\frac{\pi}{2} - \frac{\pi}{180}\right)$, $\mathbb{P}(R \leq r) = 0$.

• Si $r > \frac{v^2}{g}$, $\mathbb{P}(R \leq r) = 1$.

Ceci nous donne la fonction de répartition, donc la loi de R .

Fiche 81

81.1 Les douze forages étant indépendants, et la probabilité de trouver du pétrole dans chaque étant $\frac{1}{4}$, X suit la loi binomiale de paramètres 12 et $\frac{1}{4}$. En moyenne, $\mathbb{E}(X) = 3$ forages donneront du pétrole.

81.2 Soit X le nombre de jouets pêchés par l'enfant. X suit la loi hypergéométrique de paramètres $N = 100$; $p = 0,25$; $n = 10$.

$$1. \mathbb{P}(X = 5) = \frac{(75 \times \dots \times 71) \times (25 \times \dots \times 21) \times (10 \times \dots \times 6)}{(100 \times \dots \times 91) \times 5!} = 1,710^{-6}.$$

$$2. \mathbb{P}(X = 1) = \frac{25 \times (75 \times \dots \times 67) \times 10}{(100 \times \dots \times 91)} = 0,18.$$

81.3

$\mathbb{E}(X) = 0 \times \mathbb{P}(X = 0) + 1 \times \mathbb{P}(X = 1) = p$; $\text{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \mathbb{E}(X) - \mathbb{E}(X)^2 = p(1-p)$ car $X^2 = X$ puisque X est à valeurs dans $\{0; 1\}$. Pour tout $p \in [0; 1]$, $\text{Var}(X) = p(1-p) \leq \frac{1}{4}$ et $\text{Var}(X) = \frac{1}{4}$ si $p = \frac{1}{2}$.

81.4

X suit la loi uniforme sur $\{1; \dots; 6\}$; $\mathbb{E}(X) = \sum_{k=1}^6 k \times \frac{1}{6} = \frac{7}{2}$; $\text{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \sum_{k=1}^6 k^2 \times \frac{1}{6} - \left(\frac{7}{2}\right)^2 = \frac{6^2 - 1}{12} \approx 2,92$.

81.5 Les n lancers sont indépendants, et la probabilité de tomber sur 6 est $\frac{1}{6}$, X suit la loi binomiale de paramètres n et $\frac{1}{6}$; $\mathbb{E}(X) = \frac{n}{6}$; $\text{Var}(X) = n \times \frac{1}{6} \times (1 - \frac{1}{6}) = \frac{5n}{36}$.

81.6 $\mathbb{E}(X) = a$, $\text{Var}(X) = 0$.

Fiche 82

$$82.1 \quad \mathbb{E}(X) = \sum_{k=0}^{+\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \lambda \sum_{k=1}^{+\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda; \quad \mathbb{E}(X(X-1)) = \sum_{k=0}^{+\infty} k(k-1) \frac{\lambda^k}{k!} e^{-\lambda} = \lambda^2, \text{ et } \text{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \mathbb{E}(X(X-1)) + \mathbb{E}(X) - \mathbb{E}(X)^2 = \lambda.$$

82.2

1. Le nombre moyen d'éruptions sur l'intervalle de temps $[0; t]$ est $\mathbb{E}(X_t) = \lambda t$.

2. Pour $t = 100$, nous avons $\mathbb{E}(X_{100}) = 100$, donc $\lambda = 1$.

3. La probabilité d'observer au moins une éruption sur les cinq ans à venir est $\mathbb{P}(X_5 \geq 1) = 1 - \mathbb{P}(X_5 = 0) = 1 - e^{-5\lambda} \approx 0,993$; sur le siècle à venir, cette probabilité est $\mathbb{P}(X_{100} \geq 1) = 1 - e^{-100\lambda} \approx 1$.

4. Le nombre d'éruptions le plus probable en un an est le mode de X_1 , c'est-à-dire 0 ou 1 car $\mathbb{P}(X_1 = 0) = \mathbb{P}(X_1 = 1) > \mathbb{P}(X_1 = k)$ pour tout $k > 1$.

82.3

1. Il y a $\mathbb{E}(X) = 200 \times 10^3 \mu = 8$ défauts en moyenne sur 200 km de caténaire.

2. Le nombre de défauts le plus probable sur cette longueur est le mode de X , qui vaut ici 7 ou 8 car $\mathbb{P}(X = 7) = \mathbb{P}(X = 8) \approx 0,14 > \mathbb{P}(X = k)$ pour tout $k \in \mathbb{N} \setminus \{7; 8\}$.

3. La probabilité de n'avoir aucun défaut sur le câble est $\mathbb{P}(X = 0) = e^{-8} = 3,310^{-4}$.

82.4 Soit, pour chaque mine, p la probabilité de trouver de l'or et/ou de l'argent dans la mine : $p = \mathbb{P}(Or \cup Diamant) = \mathbb{P}(Or) + \mathbb{P}(Diamant) - \mathbb{P}(Or \cap Diamant) = 0,13$. Le nombre X de mines à parcourir avant de trouver de l'or et/ou du diamant suit une loi géométrique de paramètre p . Le nombre Y de mines à parcourir avant de trouver de l'or suit une loi géométrique de paramètre $p' = 0,10$.

82.5 Soit X la durée de vie d'un appareil ménager, avant qu'il ne tombe en panne. X suit la loi géométrique de paramètre p . Nous cherchons t tel que $\mathbb{P}(X \leq t) \leq 0,25$: la durée de garantie t doit être inférieure ou égale au quantile d'ordre 0,25 de la loi de X .

82.6 Soit X le nombre d'œufs pondus par la poule avant d'obtenir ses trois poussins. X suit une loi binomiale négative de paramètres 3 et $\frac{1}{4}$. Le nombre moyen d'œufs que la poule doit pondre avant d'avoir ses trois poussins est $\mathbb{E}(X) = 3 \times 4 = 12$. La probabilité que la poule dusse pondre 10 œufs pour avoir ses trois poussins est $\mathbb{P}(X = 10) = \frac{9 \times 8}{2} \times \frac{1}{4^3} \times \frac{3^7}{4^7} \approx 0,075$.

Fiche 83

83.1 Nous utilisons la table de la loi $\mathcal{N}(0; 1)$.

1. **2.** $z_1 = 1,960$; $z_2 = -1,960$; $z_3 = 1,645$.
 $\mathbb{P}(Z \leq 1,64) = 0,9495$; $\mathbb{P}(Z \leq -1,64) = 1 - \mathbb{P}(Z \leq 1,64) = 0,0505$.

3. Soit Φ la fonction de répartition de la loi normale centrée et réduite. $\mathbb{P}(|Z| < 1) = 2 \times \Phi(1) - 1 = 0,6826$; $\mathbb{P}(-2 < Z < 1) = \Phi(1) + \Phi(2) - 1 = 0,8185$; $\mathbb{P}(|Z - 1| > 0,5) = 1 - \Phi(1,5) + \Phi(0,5) = 0,7583$.

4. $\mathbb{E}(Z) = \int_{\mathbb{R}} z f_Z(z) dz = 0$ et $\text{Var}(Z) = \mathbb{E}(Z^2) - \mathbb{E}(Z)^2 = 1$.

83.2

Nous utilisons la table de la loi $\mathcal{N}(0; 1)$.

1. $0,975 = \mathbb{P}(X \leq x_1) = \mathbb{P}\left(\frac{X - \mu}{\sigma} \leq \frac{x_1 - \mu}{\sigma}\right)$ donc $x_1 = 1,96 \times \sigma + \mu = 4,92$. De même, $x_2 = -1,96 \times \sigma + \mu = -2,92$. Enfin, $0,90 = \mathbb{P}\left(-\frac{x_3}{\sigma} \leq \frac{X - \mu}{\sigma} \leq \frac{x_3}{\sigma}\right)$, d'où $x_3 \approx 1,64 \times \sigma = 3,28$.

2. Si σ tend vers 0, l'intervalle $[\pm x_3] \approx [\pm 1,64 \times \sigma]$ devient de plus en plus étroit autour de 0.

3. Soit Φ la fonction de répartition de la loi normale centrée et réduite. $\mathbb{P}(X \leq 4,28) = \mathbb{P}\left(\frac{X - \mu}{\sigma} \leq \frac{4,28 - \mu}{\sigma}\right) = \Phi(1,64) = 0,9495$;
 $\mathbb{P}(2 < X < 3) = \mathbb{P}\left(\frac{2 - \mu}{\sigma} < \frac{X - \mu}{\sigma} < \frac{3 - \mu}{\sigma}\right) = \Phi(1) - \Phi\left(\frac{1}{2}\right) = 0,1498$; $\mathbb{P}(X > 0) = \mathbb{P}\left(\frac{X - \mu}{\sigma} > \frac{0 - \mu}{\sigma}\right) = \Phi\left(\frac{1}{2}\right) = 0,6915$ et
 $\mathbb{P}(|X - 3| > 1) = \mathbb{P}(X > 4) + \mathbb{P}(X < 2) = 1 - \Phi\left(\frac{3}{2}\right) + \Phi\left(\frac{1}{2}\right) = 0,7583$.

4. $Z = \frac{X - \mu}{\sigma} \sim \mathcal{N}(0; 1)$ donc $\mathbb{E}(X) = \mathbb{E}(\sigma Z + \mu) = \mu$ et $\text{Var}(\sigma Z + \mu) = \sigma^2$.

83.3

Nous utilisons la table de la loi $\mathcal{N}(0; 1)$. L'intervalle de probabilité de niveau 95 % de ce que je peux recevoir avec la première action est $[100(1 + \mu_1) \pm 1,96 \times 100\sigma_1] = [22; 218]$ euros. Avec la seconde action, l'intervalle est

$[100(1 + \mu_2) \pm 1,96 \times 100\sigma_2] = [-16; 376]$ euros. Il vaut mieux choisir d'investir dans la première action.

Fiche 84

84.1

- Si X suit la loi uniforme $\mathcal{U}[a; b]$ alors $\mathbb{E}(X) = \frac{a+b}{2}$ et $\text{Var}(X) = \frac{(b-a)^2}{12}$.
- Si X suit la loi exponentielle $\mathcal{E}(\lambda)$ alors $\mathbb{E}(X) = \frac{1}{\lambda}$ et $\text{Var}(X) = \frac{1}{\lambda^2}$.
- Si X suit la loi gamma $\Gamma(a; b)$ alors $\mathbb{E}(X) = ab$ et $\text{Var}(X) = ab^2$.
- Si X suit la loi beta $\text{Beta}(a; b)$ alors $\mathbb{E}(X) = \frac{a}{a+b}$ et $\text{Var}(X) = \frac{ab}{(a+b)^2(a+b+1)}$.

84.2 $\forall k \in \mathbb{N}^*, T_k = \sum_{i=1}^k S_i, \forall i \leq k$,
 $\mathbb{E}(S_i) = \frac{1}{\lambda}$ et $\mathbb{E}(T_k) = \frac{k}{\lambda}$.

Fiche 85

85.1

1., 2. et 3. $x_1 = 2,0879$, $x_2 = 16,9190$ et $\mathbb{P}(X \geq 4,1682) = 0,9$.

85.2

1., 2., 3., 4. et 5. $t_1 = 2,1314$; $t_2 = 1,3406$; $t_3 = -1,7531$; $\mathbb{P}(T \geq 2,1314) = 0,025$ et $\mathbb{P}(T \leq -2,1314) = \mathbb{P}(T \geq 2,1314) = 0,025$.

85.3

1., 2. et 3. $0,05 = \mathbb{P}(F \leq f_1) = \mathbb{P}\left(\frac{1}{F} \geq \frac{1}{f_1}\right)$ donc $f_1 = \frac{1}{4,5337} = 0,2206$; $f_2 = 6,1631$; $\mathbb{P}(F \geq 9,1973) = 0,025$.

85.4 Soient les variables aléatoires X de loi $\chi^2(n)$ et Y de loi $\Gamma\left(\frac{n}{2}; 2\right)$. X et Y ont les mêmes densités, donc elles ont les mêmes lois.

85.5

Soient X de loi de Cauchy et Y de loi $t(1)$. X et Y ont les mêmes densités, donc elles ont les mêmes lois.

85.6

Soient T de loi $t(n)$, F_T sa fonction de répartition et f_T sa densité. Nous calculons la fonction de répartition de $-T$, notée F_{-T} , puis la densité de $-T$, notée $f_{-T} = \frac{dF_{-T}}{dt}$. Nous remarquons que T et $-T$ ont les mêmes densités, donc elles ont les mêmes lois.

Ensuite, $\forall x \in \mathbb{R}_+$, $F_{T^2}(x) = \mathbb{P}(T^2 \leq x) = \mathbb{P}(-\sqrt{x} \leq T \leq \sqrt{x}) = F_T(\sqrt{x}) - F_T(-\sqrt{x})$. La densité de T^2 est la dérivée de F_{T^2} : $f_{T^2}(x) =$

$$\frac{f_T(\sqrt{x}) + f_T(-\sqrt{x})}{2\sqrt{x}} = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\Gamma\left(\frac{n}{2}\right)} \times \frac{1}{\sqrt{x}} \times$$

$\left(1 + \frac{x}{n}\right)^{-\frac{n+1}{2}} \times \mathbb{1}_{x>0}$. Nous remarquons que T^2 suit une loi de Fisher $F(1; n)$.

Fiche 86

86.1

1. La loi de (X, Y) est donnée par : $X(\Omega) \times Y(\Omega) = \{0; 1\} \times \{1; 2; 3\}$ et $p_{X,Y}(0; 1) = \frac{0,20}{3}$; $p_{X,Y}(1; 1) = \frac{0,80}{3}$; $p_{X,Y}(0; 2) = \frac{0,80}{3}$; $p_{X,Y}(1; 2) = \frac{0,20}{3}$; $p_{X,Y}(0; 3) = \frac{0,50}{3}$; $p_{X,Y}(1; 3) = \frac{0,50}{3}$.

2. Les lois marginales sont $p_X(0) = p_X(1) = 0,50$ et $p_Y(1) = p_Y(2) = p_Y(3) = \frac{1}{3}$.

3. Pour une étudiante du campus, la probabilité qu'elle soit de l'école 2 est $\mathbb{P}(Y = 2 | X = 1) = \frac{2}{15}$.

86.2

1. La loi de (X, Y) est donnée par : $X(\Omega) \times Y(\Omega) = \{(0; 0), (0; 1), (1; 1), (1; 0)\}$ et $p_{X,Y}(0; 0) = \frac{\ell+v}{N} \times \frac{\ell+v-1}{N-1}$; $p_{X,Y}(0; 1) =$

$$\frac{\ell+v}{N} \times \frac{b}{N-1} ; p_{X,Y}(1; 1) = \frac{b}{N} \times \frac{b-1}{N-1} ; p_{X,Y}(1; 0) = \frac{b}{N} \times \frac{\ell+v}{N-1}.$$

2. Les lois marginales sont $p_X(0) = \frac{\ell+v}{N}$; $p_X(1) = \frac{b}{N}$ et $p_Y(0) = \frac{\ell+v}{N}$; $p_Y(1) = \frac{b}{N}$. X et Y ont les mêmes lois.

3. La loi de $Z = XY$ est une loi de Bernoulli de paramètre $p = \frac{b(b-1)}{N(N-1)}$.

86.3

1. $a = 1$.

2. X et Y ont toutes les deux la même loi : la loi géométrique de paramètre $p = \frac{1}{2}$.

Fiche 87

$$\mathbf{87.1} \quad \mathbb{P}(a < X \leq b, c < Y \leq d) = \int_a^b \int_c^d f(x, y) dx dy.$$

87.2 Les lois marginales sont égales, et définies par $\forall x \in \mathbb{R}$, $f_X(x) = f_Y(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$.

87.3 La densité de (X, Y) est définie par $\forall x, y \in \mathbb{R}$, $f_{X,Y}(x, y) = \mathbb{1}_{x \in [0; 1], y \in [0; 1]}$. Les densités marginales sont définies par $\forall x \in \mathbb{R}$, $f_X(x) = f_Y(x) = \mathbb{1}_{x \in [0; 1]}$.

87.5

1. $C = 2$.

2. $\forall x \in \mathbb{R}$, $f_X(x) = 2e^{-2x} \mathbb{1}_{x \geq 0}$: X suit la loi $\mathcal{E}(2)$. De même, Y suit la loi $\mathcal{E}(1)$.

3.a. $\mathbb{P}(X < 1, Y > 1) = (1 - e^{-2}) \times e^{-1}$,

b. $\mathbb{P}(X > Y) = \int_0^{+\infty} \int_y^{+\infty} f(x, y) dx dy = \frac{1}{3}$,

c. $\mathbb{P}(X < 1) = 1 - e^{-2}$.

Fiche 88

$$\mathbf{88.1} \quad \mathbb{P}(a < X \leq b, c < Y \leq d) = F_{X,Y}(b, d) - F_{X,Y}(b, c) - F_{X,Y}(a, d) + F_{X,Y}(a, c).$$

88.2

1. $\forall u \in \mathbb{R}^+$, $F_U(u) = F_{X,Y}(u, u)$.

$\forall v \in \mathbb{R}^+, F_V(v) = \mathbb{P}(X \leq v \text{ ou } Y \leq v) = F_{X,Y}(v, +\infty) + F_{X,Y}(+\infty, v) - F_{X,Y}(v, v).$

2. $\forall u, v \in \mathbb{R}^+, F_{U,V}(u, v) = F_{X,Y}(\min(u, v), u) + F_{X,Y}(u, \min(u, v)) - F_{X,Y}(\min(u, v), \min(u, v)).$ Le calcul de $F_{U,V}(u, +\infty)$ et $F_{U,V}(+\infty, v)$ donne bien $F_U(u)$ et $F_V(v)$ calculées précédemment.

3. $\mathbb{P}(U > 300, V \leq 100) = F_V(100) - F_{U,V}(300; 100).$

Fiche 89

89.1 Le temps moyen d'attente est $\mathbb{E}(T)$ où $T = |X - Y|$ est le temps d'attente. $\mathbb{E}(T) = \int_{-1}^1 \int_{-1}^1 |x - y| dx dy = \frac{1}{3}.$

89.2 Pour $i = 1; \dots; 7$, soit $X_i = 1$ si la $i^{\text{ème}}$ famille est représentée dans les n paquets, et $X_i = 0$ sinon. Pour tout $i = 1; \dots; 7$, X_i suit une loi de Bernoulli de paramètre $p = 1 - (1 - \frac{1}{7})^n = \mathbb{P}(X_i = 1)$. Le nombre moyen de figurines représentées est $\mathbb{E}(X_1 + \dots + X_7) = 7 \times p.$

89.3

1. $\mathbb{E}(X) = \mathbb{E}(Y) = 0$,
et $\mathbb{E}(XY) = \int_{\mathbb{R}} \int_{\mathbb{R}} xy f_{X,Y}(x, y) dx dy = \frac{1}{\pi} \int_{-1}^1 \left(\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} xy dy \right) dx = 0$, d'où $\rho(X, Y) = 0$.

2. $\text{Cov}(R, S) = a \mathbb{V}\text{ar}(S), \sqrt{\mathbb{V}\text{ar}(R)} = |a| \sqrt{\mathbb{V}\text{ar}(S)}$ d'où $\rho(R, S) = \frac{a \mathbb{V}\text{ar}(S)}{|a| \mathbb{V}\text{ar}(S)} = 1$ si $a > 0$ et -1 si $a < 0$.

89.4

- Utiliser la linéarité de l'espérance.
- Évident.
- Évident.
- Utiliser la linéarité de l'espérance, dans la définition de la covariance.
- Utiliser la linéarité de l'espérance, dans la définition de la covariance.

Fiche 90

90.1

1. $\forall y \in \mathbb{R}, F_Y(y) = 1 - (1 - F_X(y))^n$; et $F_Z(y) = (F_X(y))^n.$

2. $\forall x \in \mathbb{R}, F_X(x) = (1 - e^{-\lambda x}) \mathbb{1}_{x \geq 0}$, donc $\forall y \in \mathbb{R}, F_Y(y) = (1 - e^{-\lambda ny}) \mathbb{1}_{y \geq 0}$: Y suit la loi $\mathcal{E}(n\lambda)$; et $F_Z(y) = (1 - e^{-\lambda y})^n \mathbb{1}_{y \geq 0}.$

90.2 Dans l'Exercice 89.3, nous montrons que $\rho(X, Y) = 0$. Or, le couple aléatoire (X, Y) représente les coordonnées d'un point jeté au hasard dans le disque unité, donc il vérifie $X^2 + Y^2 \leq 1$: X et Y ne sont donc pas indépendantes.

Fiche 91

91.1 $\forall u \in \mathbb{R}^+, F_U(u) = \mathbb{P}(X \leq uY) = \int_0^{+\infty} \left(\int_0^{uy} 2e^{-2x} dx \right) e^{-y} dy = \frac{2u}{1+2u}$, et $F_U(u) = 0$ si $u < 0$. La densité de U est donc $f_U(u) = F'_U(u) = \frac{2}{(1+2u)^2} \mathbb{1}_{u \geq 0}.$

91.3 $(E, V) = h(M, V)$ où la fonction h est définie par

$$h : \begin{cases} (\mathbb{R}_+^*)^2 \longrightarrow (\mathbb{R}_+^*)^2 \\ (x, y) \longmapsto (\frac{1}{2}xy^2, y) \end{cases}$$

h est un $\mathcal{C}^1$ -difféomorphisme sur $(\mathbb{R}_+^*)^2$: $h^{-1}(u, v) = (\frac{2u}{v^2}, v).$

La densité de (E, V) est $\forall (u, v) \in \mathbb{R}_+^2$,
 $f_{E,V}(u, v) = f_{M,V}\left(\frac{2u}{v^2}, v\right) \times \left| \frac{2}{v^2} - \frac{4u}{v^3} \right| \times \mathbb{1}_{u>0, v>0} = f_M\left(\frac{2u}{v^2}\right) f_V(v) \times \frac{2}{v^2} \times \mathbb{1}_{u>0, v>0}$ et la densité de E est $f_E(u) = \int_0^{+\infty} f_M\left(\frac{2u}{v^2}\right) f_V(v) \times \frac{2}{v^2} dv \times \mathbb{1}_{u>0}.$

L'énergie moyenne est $\mathbb{E}(E) = \frac{1}{2} \mathbb{E}(M) \mathbb{E}(V^2) = \frac{1}{2} \mathbb{E}(M) (\mathbb{V}\text{ar}(V) + \mathbb{E}(V)^2).$

91.4 Méthode 1 :

$\forall x \in \mathbb{R}^+, \mathbb{P}(X_2 \leq x) = \mathbb{P}(S_1 + S_2 \leq x) = \int_0^{+\infty} \int_0^{+\infty} \lambda^2 e^{-\lambda(u+v)} \mathbb{1}_{u+v \leq x} dudv =$

$(1 - e^{-\lambda x} - \lambda x e^{-\lambda x}) \mathbb{1}_{x \geq 0}$, puis $f_{X_2}(x) = \lambda^2 x e^{-\lambda x} \mathbb{1}_{x \geq 0}$. La loi de X_2 est la loi $\Gamma(2; 1/\lambda)$.

Méthode 2 : chercher la loi de (X_1, X_2) , puis en déduire la loi de X_2 .

91.5

1. La loi de (U, V) est définie par $\forall (u, v) \in \mathbb{R}^2$, $F_{U,V}(u, v) = \begin{cases} F_X(u)^2 - (F_X(u) - F_X(v))^2 & \text{si } u \geq v \\ F_X(u)^2 & \text{si } u < v \end{cases}$ où $F_{U,V}$ et F_X sont les fonctions de répartition de (U, V) et de X .

La densité $f_{U,V}$ est définie par $f_{U,V}(u, v) = \frac{\partial^2 F_{U,V}(u, v)}{\partial u \partial v} = 2f(v)f(u) \mathbb{1}_{u \geq v}$.

2. U et V sont dépendantes car $U \geq V$.

91.6

1. $M(\Omega) = \mathbb{N}$, et $\forall k \in \mathbb{N}$, $\mathbb{P}(M = k) = \sum_{n=k}^{+\infty} \mathbb{P}(M = k | N = n) \mathbb{P}(N = n) = e^{-\lambda p} \frac{(\lambda p)^k}{k!}$. M suit la loi de Poisson de paramètre λp .

2. et 3. $\mathbb{E}(M) = \text{Var}(M) = \lambda p$.

Fiche 92

92.1 Soit X la hauteur de pluie journalière, telle que $\mathbb{E}(X) = \mu$ alors $\mathbb{P}(X \geq 30) \leq \frac{\mu}{30}$.

92.2 $\mathbb{P}(E \leq \varepsilon) = \mathbb{P}(|X - \mu| \leq \mu \varepsilon) \geq 1 - \frac{\text{Var}(X)}{(\mu \varepsilon)^2} = 1 - \frac{1}{r^2 \varepsilon^2}$.

92.3

1.a. L'intervalle de probabilité est $[\mu \pm a] = [492, 2; 1207, 8]$.

b. Il est possible que cette planche provienne du matériau $\mathcal{M}$.

2.a. La loi des grands nombres donne une approximation de $\mathbb{E}(X)$ par $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$.

b. $n = 64 \times 10^9$.

92.4

1. Z_1 suit une loi de Bernoulli de paramètre $p = \frac{\pi}{4}$. $\mathbb{E}(Z_1) = p$, $\text{Var}(Z_1) = p(1-p) \leq \frac{1}{4}$. La suite des $(Z_i)_{i \geq 1}$ est une suite de variables aléatoires indépendantes et de même loi.

2. Il suffit d'appliquer la loi faible des grands nombres et d'utiliser le fait que $\text{Var} Z_1 \leq \frac{1}{4}$.

3. D'après la loi des grands nombres, $\bar{Z}_{1000}$ approche $\mathbb{E}(Z_1) = \frac{\pi}{4}$, une approximation de π est donc $4\bar{Z}_{1000}$, c'est-à-dire 3,144. Un intervalle de confiance pour π , de niveau supérieur ou égal à 95 % est donné par $\left[4\bar{Z}_{1000} \pm 4 \sqrt{\frac{10^{-2}}{2}} \right] = [2, 861; 3, 427]$.

4. Cherchons n tel que $0,95 = \mathbb{P}(|4\bar{Z}_n - \pi| \leq 4a) \geq 1 - \frac{1}{4na^2}$ avec $4a = 10^{-2}$. Nous obtenons $n = 8 \times 10^5$.

Fiche 93

93.1 Nous appliquons le Théorème de la limite centrale. La probabilité de recevoir moins de 100 appels en 2 heures est $\mathbb{P}(S_{100} \geq 120) = \mathbb{P}\left(\frac{\frac{1}{100} \sum_{i=1}^{100} X_i - \frac{1}{\lambda}}{\frac{1}{\lambda}/\sqrt{100}} \geq \frac{\frac{120}{100} - \frac{1}{\lambda}}{\frac{1}{\lambda}/\sqrt{100}}\right) \approx \mathbb{P}(Z \geq 2) = 0,0228$ où $Z \sim \mathcal{N}(0; 1)$.

93.2 Les variables aléatoires P_i sont supposées indépendantes et de même loi. Nous appliquons le Théorème de la limite centrale, et cherchons n tel que $0,05 \geq \mathbb{P}(\sum_{i=1}^n P_i \geq 600) = \mathbb{P}\left(\frac{\frac{1}{n} \sum_{i=1}^n P_i - \mu}{\sigma/\sqrt{n}} \geq \frac{\frac{600}{n} - \mu}{\sigma/\sqrt{n}}\right) \approx \mathbb{P}\left(Z \geq \frac{\frac{600}{n} - \mu}{\sigma/\sqrt{n}}\right)$ où $Z \sim \mathcal{N}(0; 1)$, d'où il faut $\frac{\frac{600}{n} - \mu}{\sigma/\sqrt{n}} \geq 1,645$. Le plus grand n vérifiant cette inégalité est $n = 9$.

93.3 La probabilité que l'action vaille plus de 2 euros dans un an est $\mathbb{P}(Y_{365} \geq 2) = \mathbb{P}\left(\frac{\frac{1}{365} \sum_{i=1}^{365} X_i}{\sigma/\sqrt{365}} \geq \frac{1}{\sqrt{365}}\right) \approx \mathbb{P}\left(Z \geq \frac{1}{\sqrt{365}}\right) \approx 0,48$ où $Z \sim \mathcal{N}(0; 1)$. Nous avons appliqué le Théorème de la limite centrale.

93.5

1. $v = \sum_{i=1}^4 p_i v_i$. Notons $X_i, i = 1; \dots; 500$ le prix de vente du $i^{\text{ème}}$ œuf. Les variables aléatoires X_i sont indépendantes et de même loi. Nous avons $\forall i, \mathbb{E}(X_i) = v$.

2. La probabilité que le fermier gagne plus de 500v euros avec la vente des 500 œufs du jour est $\mathbb{P}\left(\sum_{i=1}^{500} X_i \geq 500v\right) =$

$$\mathbb{P}\left(\frac{\frac{1}{500} \sum_{i=1}^{500} X_i - v}{\sqrt{\text{Var}(X_1)/500}} \geq 0\right) \approx \mathbb{P}(Z \geq 0) = 0,5$$

où $Z \sim \mathcal{N}(0; 1)$. Nous avons appliqué le Théorème de la limite centrale.

Fiche 94

94.1

1. Soit X le nombre d'œufs cassés ce jour là, X suit la loi binomiale de paramètres 500 et p .

2. La probabilité qu'il n'y ait pas plus de 20 œufs cassés ce jour là est $\mathbb{P}(X \leq 20) \approx \mathbb{P}(Y \leq 20) = e^{-25} \left(1 + \frac{25}{1!} + \frac{25^2}{2!} + \dots + \frac{25^{20}}{20!}\right) = 0,185$ où Y suit une loi de Poisson de paramètre 25.

94.2 Soit X le nombre de personnes atteinte de la grippe saisonnière en France, X suit la loi binomiale de paramètres $n = 65$ millions et $p = 0,45$. La probabilité qu'au moins 29 258 022 personnes soient malades est approximativement $\mathbb{P}(Y \geq 29\,258\,022)$ où $Y \sim \mathcal{N}(np, \sqrt{np(1-p)})$. Nous avons $\mathbb{P}(Y \geq 29\,258\,022) = \mathbb{P}(Z \geq \frac{29\,258\,022 - np}{\sqrt{np(1-p)}}) = 0,0228$, où $Z \sim \mathcal{N}(0; 1)$.

94.3 La probabilité que plus de 3100 personnes passent par la gare de Paris-Est lors d'un week-end de l'Ascension est $\mathbb{P}(X \geq 3100) = \mathbb{P}\left(\frac{X - 3000}{\sqrt{3000}} \geq \frac{100}{\sqrt{3000}}\right) \approx \mathbb{P}(Z \geq 1,82) = 0,0344$, où $Z \sim \mathcal{N}(0; 1)$.

94.4 Soit X le nombre de brochets prélevés parmi les cinq poissons. X suit une loi hypergéométrique de paramètres $N; p; n$. La probabilité d'obtenir 2 brochets est $\mathbb{P}(X = 2) \approx \mathbb{P}(Y =$

$2) = \binom{5}{2} p^2 (1-p)^3 = 0,07$, où Y suit la loi binomiale de paramètres $n = 5; p = 0,1$.

Fiche 96

96.1 Pour la première application de la Fiche 97, la population étudiée est l'ensemble des 27 États membres de l'U.E. Les individus sont les 27 États membres. Le caractère étudié est le nombre de personnes vivant dans chaque État membre, qui est une variable quantitative discrète.

Pour la seconde application de la Fiche 97, la population étudiée est l'ensemble des ingénieurs français de moins de 65 ans en 2010. La variable étudiée est l'âge de l'ingénieur, qui est une variable quantitative continue.

Pour la première application de la Fiche 98, la population étudiée est l'ensemble des ménages français de moins de 6 personnes en 2008. Le caractère étudié est le nombre de personnes constituant le ménage. Il varie de 1 à 6 et c'est une variable quantitative discrète.

Pour les autres séries statistiques, voir les suppléments en ligne.

Fiche 97

97.1 Pour la première application de la Fiche 98,

Nombre de personnes	Effectifs des ménages	Effectifs cumulés
1	8 914 849	8 914 849
2	8 776 420	17 691 269
3	3 911 419	21 602 688
4	3 303 676	24 906 364
5	1 240 991	26 147 355
6	468 123	26 615 478

Pour la seconde application de la Fiche 98,

Âge	Effectifs	Effectifs cumulés
20 à 30 ans	187 850	187 850
30 à 35 ans	130 050	317 900
35 à 40 ans	115 600	433 500
40 à 45 ans	86 700	520 200
45 à 50 ans	65 025	585 225
50 à 55 ans	57 800	643 025
55 à 60 ans	43 350	686 375
60 à 65 ans	36 125	722 500

Pour les autres séries statistiques, voir les suppléments en ligne.

Fiche 98

98.1 Voir les suppléments en ligne pour les représentations graphiques.

Fiche 99

99.1 Voir les suppléments en ligne pour les trois représentations graphiques demandées.

Fiche 100

100.1

1. et 2. Voici les valeurs des caractéristiques demandées :

	μ	Q_1	Q_2	Q_3
Partiel Ens.	15,64	8,30	12,50	24,00
Partiel Hom.	8,19	5,40	7,10	10,00
Partiel Fem.	24,66	11,40	19,60	39,00
Non-Salariés	47,55	45,50	47,90	49,70

100.2

1. La moyenne est $\mu(X) = 2,27$.

2. Tableau des fréquences cumulées.

Nbre de Personnes	Fréquences cumulées
1	0,33
2	0,66
3	0,81
4	0,94
5	0,98
6	1,00

Ainsi $Q_1(X) = 1$, $Q_2(X) = 2$ et $Q_3(X) = 3$.

Fiche 101

101.1

1. L'écart-type est $\sigma(X) = 1,267$

2. L'exercice 100.2 donne $Q_2(X) = 2$. Les fréquences cumulées de la série $|x_i - Q_2(X)|$ permettent de déterminer que la médiane des écart-absolus à la médiane est $MAD(X) = 1$.

3. L'exercice 100.2 donne $Q_1(X) = 1$ et $Q_3(X) = 3$. Ainsi, l'étendue interquartile est $EIQ(X) = 3 - 1 = 2$.

4. Le coefficient de variation est $CV(X) = 55,82\%$.

101.2

1., 2., 3. et 4. En utilisant les caractéristiques, moyenne et quantiles, calculées à l'Exercice 100.1 :

	σ	MAD	EIQ	CV
Partiel Ens.	9,94	5,40	15,70	63,51
Partiel Hom.	4,75	2,60	4,60	57,98
Partiel Fem.	16,83	10,30	27,60	68,28
Non-Salariés	3,67	2,30	4,20	7,72

101.3

1. et 2.

	μ	Q_1	Q_2	Q_3
Partiel Ens.	15,37	7,50	14,50	21,79
Partiel Hom.	7,59	3,21	6,43	9,64
Partiel Fem.	25,00	11,94	19,44	38,12
Salariés	40,65	38,38	41,03	43,01
Non-Salariés	47,69	45,55	47,66	49,77

3., 4. et 5.

	σ	EIQ	CV (%)
Partiel Ens.	9,62	14,29	62,56
Partiel Hom.	5,16	6,43	67,94
Partiel Fem.	17,00	26,18	67,99
Salariés	2,41	4,64	5,94
Non-Salariés	3,96	4,22	8,31

6. Les valeurs ne sont pas identiques. Les caractéristiques obtenues avec les distributions groupées ne sont donc que des approximations des valeurs réelles de la population.

7. Lorsque c'est possible, il faut travailler avec les valeurs non groupées pour obtenir les valeurs exactes des caractéristiques.

Fiche 102

102.1

1. Pour la représentation graphique de la boîte à moustaches, nous renvoyons le lecteur aux suppléments en ligne.

2. et 3.

	γ_1	β_1	γ_2	β_2
Partiel Ens.	1,34	1,79	2,25	5,25
Partiel Hom.	1,72	2,96	4,02	7,02
Partiel Fem.	1,03	1,07	0,98	3,98

102.2

1. et 2.

	γ_1	β_1	γ_2	β_2
Non-Salariés	0,08	0,01	0,02	3,02

102.3

1. et 2.

	γ_1	β_1	γ_2	β_2
Partiel Ens.	0,93	0,86	0,54	3,54
Partiel Hom.	1,87	3,51	1,17	4,17
Partiel Fem.	0,91	0,82	0,37	3,37
Salariés	-0,54	0,29	-1,71	1,29
Non-Salariés	-0,07	0,00	-1,15	1,85

Fiche 103

103.1

1. Un calcul rapide des dénombrements nous permet de retrouver le tableau fourni dans l'énoncé.

2. La distribution marginale des effectifs de X est :

(35,40]	10
(40,45]	17

La distribution marginale des effectifs de Y est :

(35,40]	(40,45]	(45,50]	(50,55]	(55,60]
1	4	16	5	1

La distribution marginale des fréquences de X est :

(35,40]	0,37
(40,45]	0,63

La distribution marginale des fréquences de Y est :

(35,40]	(40,45]	(45,50]	(50,55]	(55,60]
0,04	0,15	0,59	0,19	0,04

3. Les distributions conditionnelles des fréquences de X sachant $Y = y_j$ sont regroupées dans le tableau suivant :

	(35,40]	(40,45]	(45,50]	(50,55]	(55,60]
(35,40]	0,00	0,25	0,31	0,60	1,00
(40,45]	1,00	0,75	0,69	0,40	0,00

4. Les distributions conditionnelles des fréquences de Y sachant $X = x_i$ sont regroupées dans le tableau suivant :

	(35,40]	(40,45]	(45,50]	(50,55]	(55,60]
(35,40]	0,00	0,10	0,50	0,30	0,10
(40,45]	0,06	0,18	0,65	0,12	0,00

Fiche 105

105.1

1. La covariance est égale à 248,4217. Le coefficient de corrélation linéaire est égal à 0,9661348.

2. Comme le coefficient de corrélation linéaire est très proche de 1, nous pouvons en déduire que les deux caractères sont très liés et de façon positive, c'est-à-dire qu'ils évoluent dans le même sens.

Fiche 108

108.1 Nous avons $\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$ et $\ln L = \ln \binom{n}{k} + k \ln p + (n-k) \ln(1-p)$. Nous différencions la fonction par rapport à p et nous obtenons

$$\frac{\partial \ln L}{\partial p} = \frac{k}{p} - \frac{n-k}{1-p} = 0.$$

Ainsi la seule valeur annulant la dérivée est $\frac{k}{n}$. Calculons la dérivée seconde pour vérifier que la log-vraisemblance admet un maximum local en $\frac{k}{n}$.

$$\frac{\partial^2 \ln L}{\partial p^2} = -\frac{k}{p^2} - \frac{n-k}{(1-p)^2}.$$

Et donc

$$\left(\frac{\partial^2 \ln L}{\partial p^2} \right)_{p=k/n} = -\frac{n^3}{k(n-k)} < 0.$$

Remarquons que lorsque p tend vers 0 ou quand p tend vers 1, $\ln L$ tend vers $-\infty$. Le maximum local précédent est donc l'unique maximum global. L'estimateur par maximum de vraisemblance est ainsi $\hat{p}_n = K/n$.

108.2 Les estimateurs par maximum de vraisemblance et par la méthode des moments sont égaux à $\hat{\mu}_n = \frac{1}{n} \sum_{i=1}^n X_i$.

Fiche 109

109.1

1. La plus petite valeur de l'échantillon qui est 4 est une estimation du minimum de la population et la plus grande valeur de l'échantillon

qui est 17 est une estimation du maximum de la population.

2. L'estimation de la moyenne de la population est égale à 9,56.

3. L'estimation de la médiane de la population est égale à 9,00.

4. L'estimation du premier quantile de la population est égale à 7,00 et l'estimation du troisième quantile de la population est égale à 11,25.

Fiche 110

110.1

1. Une estimation de l'étendue est égale à $17 - 4 = 13$.

2. L'intervalle interquantile de l'échantillon est égal à $[7, 00; 11, 25]$. Une estimation de l'étendue interquantile est égale à $11, 25 - 7, 00 = 4, 25$.

3. Une estimation biaisée de la variance est égale à 9, 3664. Une estimation sans biais de la variance est égale à 9, 557551. Une estimation biaisée de l'écart-type est égale à 3, 091529. Une estimation sans biais de l'écart-type est égale à 3, 107341.

4. Une estimation du coefficient de variation est égale à 0, 3233817.

Fiche 111

111.1 Compte-tenu des résultats observés le client doit renvoyer le lot au fournisseur.

Fiche 112

112.1

1. L'intervalle de confiance de la moyenne des 200 copies est :

$$\left[11 - 1,960 \times \frac{2}{\sqrt{7}}; 11 + 1,960 \times \frac{2}{\sqrt{7}} \right] \\ \approx]9,52; 12,48[.$$

2. Si l'amplitude de l'intervalle de confiance est égale à 2, nous devons avoir :

$$1,96 \times \frac{2}{\sqrt{n}} = 1,$$

ce qui donne $n \approx 15,4$. En corrigeant 16 copies, l'enseignant peut situer la moyenne de ses étudiants.

3. Il faut que l'intervalle de confiance à 99 % soit égal à $]10; 12[$. Nous devons donc avoir $2,576 \times \frac{2}{\sqrt{n}} = 1$, ce qui donne $n \approx 26,5$. Si l'enseignant corrige 27 copies et qu'il trouve une moyenne égale à 11, il peut dire que la moyenne de ses étudiants est supérieure à 10, avec un risque d'erreur de 1 %.

Fiche 114

114.1

1. La statistique du test est égale à $\chi_{n-1,c}^2(obs) = (99 \cdot 9)/8 = 111,375$. La zone d'acceptation est $]73,36; 128,42[$. Comme la statistique du test appartient à la zone d'acceptation, le test n'est pas significatif. Nous conservons $\mathcal{H}_0$ avec un risque de deuxième espèce β qu'il faudrait évaluer.

2. La statistique du test est égale à $Z_n(obs) = \frac{4-3}{\sqrt{8}/\sqrt{100}} = \frac{5\sqrt{2}}{2} \approx 3,536$.

La zone d'acceptation est $] - 1,960; 1,960[$. Comme la statistique du test n'appartient pas à la zone d'acceptation, le test est significatif. Nous rejetons $\mathcal{H}_0$ et décidons que $\mathcal{H}_1$ est vraie avec un risque d'erreur de première espèce α égal à 5%.

3. La statistique du test est égale à $T_{n-1}(obs) = \frac{4-3}{3/\sqrt{100}} = \frac{10}{3} \approx 3,333$.

Pour un risque α fixé, les valeurs critiques positives des lois de Student décroissent lorsque le nombre de degrés de liberté augmente. Par conséquent, nous avons : $0 < c_{99;0,975} < c_{50;0,975} = 2,0086$. La zone d'acceptation est incluse dans $] - 2,0086; 2,0086[$. Comme la statistique du test n'appartient pas à cette zone, elle ne peut appartenir à la zone d'acceptation : le test est significatif. Nous rejetons $\mathcal{H}_0$ et décidons que $\mathcal{H}_1$ est vraie avec un risque d'erreur de première espèce α égal à 5%.

4. $\widehat{\sigma}_n^2(obs) = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2 = \frac{1}{n} \sum_{i=1}^n x_i^2 - 2\widehat{\mu}_n(obs)\mu + \mu^2$ et Or $\frac{1}{n} \sum_{i=1}^n x_i^2 = \frac{n-1}{n} S_{n,c}^2 +$

$$\widehat{\mu}_n(obs)^2. \text{ Donc nous avons : } \frac{1}{n} \sum_{i=1}^n x_i^2 = \frac{9 \times 99}{100} + 16 = \frac{2491}{100} \text{ et } \widehat{\sigma}_n^2(obs) = \frac{2491}{100} - \frac{991}{100} = \frac{1500}{100} = 15.$$

$$\text{La statistique du test est égale à } \chi_n^2(obs) = \frac{100 \times \frac{991}{100}}{8} = \frac{991}{8} = 123,875.$$

La zone d'acceptation est $]77,22; 129,56[$. Comme la statistique du test appartient à la zone d'acceptation, le test n'est pas significatif. Nous conservons $\mathcal{H}_0$ avec un risque de deuxième espèce β qu'il faudrait évaluer.

114.2 Au seuil $\alpha = 5 \%$, les mesures ne sont pas compatibles avec la valeur cible car le test est significatif :

$$T_{99}(obs) = \sqrt{100} \frac{299\,852,40 - 299\,734,50}{79,01055} \approx 14,9221 \geq 1,9842 = t_{99;0,0975}.$$

Fiche 115

115.1 Le test n'est pas significatif au seuil $\alpha = 5 \%$.

$$F(obs) = \frac{6}{20} = 0,3 \in]0,2840; 3,0602[\text{ en utilisant la relation } c_{0,025;10;15} = 1/c_{0,975;15;10} = 1/3,5217.$$

Fiche 116

116.1 Les distributions sous-jacentes étant normales, nous utiliserons un test de Student, si les deux variances peuvent être supposées égales, ou un Test de Welch, si celles-ci ne peuvent pas être supposées égales. Afin de déterminer dans laquelle des deux situations nous nous trouvons, nous avons recours à un test de Fisher d'égalité des variances car les deux distributions dont nous comparons les variances sont normales. La réalisation de la statistique du test de Fisher est :

$$F = \frac{207200/9}{38200/9} = 5,4241.$$

Elle n'appartient pas à la zone d'acceptation qui est $]0,2690; 3,7168[$. Ainsi, au seuil $\alpha = 5 \%$, nous décidons que les deux variances sont différentes. Par conséquent, nous utilisons le test de Welch. La réalisation de la statistique

de test :

$$T_v(obs) = \frac{2800 - 2680}{\sqrt{\frac{20720}{10} + \frac{38200}{9}}} \approx 2,2981.$$

Pour déterminer les degrés de liberté de la loi de Student à utiliser, nous calculons :

$$\frac{\left(\frac{20720}{9} + \frac{3820}{9}\right)^2}{\frac{(20720)^2}{9^3} + \frac{(3820)^2}{9^3}} \approx 12,21.$$

L'entier le plus proche de 12,21 est 12 : la valeur critique est $c_{0,975} = 2,1788$. Le test est donc significatif au seuil de 5 % : nous rejetons l'hypothèse nulle $\mathcal{H}_0$ et décidons que l'hypothèse alternative $\mathcal{H}_1$ est vraie. Les moyennes des deux chaînes de fabrication sont significativement différentes au seuil de 5 %. Le risque associé à cette décision est un risque d'erreur de première espèce qui vaut 5 %.

Fiche 117

117.1 Nous cherchons à déterminer l'adéquation entre la loi parente de l'échantillon et une loi de Poisson. Nous souhaitons utiliser un test du khi-deux d'adéquation et commençons par remarquer que l'effectif de l'échantillon est égal à 50, ce qui est une taille suffisante pour pouvoir utiliser ce test. Afin que les autres conditions d'utilisation, portant sur les effectifs attendus, soient respectées, il est nécessaire de procéder à des regroupements en classes des valeurs prises par l'échantillon. Il y a au total 7 classes et donc $7 - 1 - 1 = 5$, car le paramètre λ de la loi de Poisson doit être estimé, degrés de liberté pour la loi du khi-deux à utiliser pour la détermination de la valeur critique. Enfin, nous trouvons $\chi^2(obs) = 3,4030 \leq 11,0705 = c_{0,95}$.

Le test n'est pas significatif au seuil $\alpha = 5 \%$: nous conservons l'hypothèse nulle $\mathcal{H}_0$ avec un risque de deuxième espèce β . Nous décidons que rien ne s'oppose à l'hypothèse que la variable « Nombre d'émissions radioactives pendant une seconde » suit une loi de Poisson. Nous pouvons faire confiance à cette décision car les conditions d'utilisation du test sont remplies.

Fiche 118

118.1 Commençons par le test de Shapiro-Wilk. En faisant les calculs à l'aide des tables, nous avons $W = 0,8562 \leq 0,881 = w_{0,05}$. Avec un logiciel utilisant la méthode Royston, comme **R** ou SAS : $W = 0,8564$ et $p\text{-value} = 0,0214$. Avec Maple : $W = 0,8598$ et $p\text{-value} = 0,0237$.

Pour le test de Shapiro-Francia, $W' = 0,8817 \leq 0,9398 = w'_{0,05}$, le test est significatif au seuil $\alpha = 5\%$.

Dans tous les cas, nous rejetons l'hypothèse nulle $\mathcal{H}_0$ et décidons que l'hypothèse alternative $\mathcal{H}_1$ est vraie avec un risque de première espèce α égal à 5 %. Nous décidons que l'hypothèse de normalité n'est pas crédible.

Index

A

abscisse curviligne 250, 270
aire d'une surface 252, 274
algorithme d'Euclide 34
algorithme du Pivot de Gauss 59
anneau 30
anneau intègre 31
antécédent 18
application 18
application linéaire
 automorphisme 78
 composée 79
 endomorphisme 78
 isomorphisme 78
 morphisme 78
 produit 79
 somme 79
approximation par une fonction
 en escalier 179
arrangements
 avec remise ou avec répétitions 292
 sans remise ou sans répétition 290
assertion 2
asymptote 175

B

base 38, 75
base orthonormée 39
bijectivité 19
bord d'un ensemble 234
borne inférieure 16, 126
borne supérieure 16, 126
boule 234
branche infinie 175, 255, 260
branche parabolique 175

C

caractère 393
 qualitatif 393
 multiple 393
 simple 393
 quantitatif 393
 multiple 393
 simple 393
 univarié 393
simple

 continu 393
 discret 393
cardinal 23
cercle asymptotique 260
cercle trigonométrique 158
champ conservatif 271
champ de vecteurs 236
champ de vecteurs à circulation conservative 276
changement de base
 pour un vecteur 91
 pour une application linéaire 91
circulation 271
circulation d'un champ de gradient 271
classe 396
classe d'équivalence 15
coefficient de corrélation 364
coefficients de Fourier complexes 210
coefficients de Fourier réels 210
cofacteur 94
colonne pivot 59
comatrice 94
combinaison 291
combinaison linéaire 74
composante 38
composantes d'un champ de vecteurs 236
composantes d'une fonction 235
congruence 35
conique
 axe focal 118
 direction 118
 ellipse 118, 119
 excentricité 118
 foyer 118
 hyperbole 118, 119, 120
 parabole 118
conjonction 3
connecteurs logiques 2
connexe par arcs 234, 271
contradiction 4
contraposée 4
convergence
 absolue 198, 203
 en loi 378
 en moyenne 218
 en probabilité 375
 normale 199, 203

presque sûre 375
 simple 194, 198
 uniforme 194, 198
 convexe 234
 coordonnées 38, 42
 corps 31
 cosinus 158
 couple aléatoire 350
 courbe paramétrée 250
 courbes de niveau 234
 courbure 251
 covariance 363

D

densité de probabilité 299, 307, 356
 densités marginales 356
 dérivée directionnelle d'une fonction 241
 dérivée partielle 238
 dérivée partielle d'ordre deux 246
 déterminant 94
 développement
 de Mac-Laurin 206
 de Taylor 207
 en série entière 206
 limité 166, 170
 limité partie régulière 170
 limité reste 170
 diagonalisation 103
 diagramme en bâtons 307
 différence 11
 différence symétrique 11
 dimension 75
 dimension finie 75
 disjonction 3
 distance 42, 115, 127
 d'un point à un plan 43
 d'un point à une droite 43
 distribution statistique
 groupée 398
 non groupée 398
 qualitative 399
 divergence 243
 diviseurs de zéro 31
 division 23, 34
 division euclidienne
 quotient 34, 47
 reste 34, 47
 domaine plan de type I 262
 domaine plan de type II 263
 données brutes 394
 droite vectorielle 76

E

écart-type 320
 échantillon 393
 effectif
 cumulé en a_i 398
 cumulé en x_i 398
 de la classe $[a_i; a_{i+1}]$ 398
 de la valeur x_i 398
 égalité de polarisation 111
 élément
 comparable 15
 de deuxième espèce 51
 de première espèce 51
 invertible 26
 régulier 27
 simple 50
 ellipse 119
 ensemble
 borné 234
 compact 234
 complémentaire 10
 convexe 234
 connexe par arcs 234
 de définition 18
 dénombrable 292
 des nombres réels 126
 disjoints 10
 équipotent 23
 fermé 234
 fini 23
 infini 24
 intersection 10
 ouvert 234
 singleton 10
 union 10
 vide 10
 entiers naturels 22
 entiers relatifs 22, 34
 équation
 aux dérivées partielles 282
 caractéristique 226
 cartésienne 42, 43
 d'un plan 43
 d'une droite 43
 de la tangente 151, 254
 différentielle 222
 différentielle linéaire d'ordre 1 222
 différentielle linéaire d'ordre 2 226
 linéaire 54
 paramétrique 43, 251, 252
 polaire 258

équiprobabilité sur un univers fini 298
 équivalence 3
 espace de probabilité 295
 espace euclidien 111
 espace préhilbertien 111
 espace propre 103
 espace vectoriel
 addition 70
 multiplication 70
 somme 71
 somme directe 71
 sous-espace vectoriel 70
 supplémentaires 71
 espaces orthogonaux 114
 espérance 314, 362
 événement 294
 élémentaire ou éventualité 294
 certain 294
 complémentaire 294
 impossible ou vide 294
 événements
 disjoints ou incompatibles 294
 indépendants 302
 mutuellement indépendants 302
 expérience aléatoire 294
 extremum global 247
 extremum local 155, 247

F

famille
 génératrice 74
 libre 75
 sommable 266, 267
 flux d'un champ vectoriel à travers
 une surface 275
 fonction 18
 analytique 206
 arccosinus 159
 arcsinus 159
 arctangente 159
 argument cosinus hyperbolique 163
 argument sinus hyperbolique 163
 argument tangente hyperbolique 163
 bornée 148
 $\mathcal{C}^1$ -difféomorphisme 242
 classe $\mathcal{C}^1$ 242
 classe $\mathcal{C}^\infty$ 152
 classe $\mathcal{C}^n$ 152
 composée 147
 concave 155
 continue 146, 235, 236

continue à droite 146
 continue à gauche 146
 continue par morceaux 179
 continue sur un intervalle 147
 convexe 155
 cosinus 158
 cosinus hyperbolique 162
 croissante 154
 de plusieurs variables 234
 de répartition 310
 de répartition conjointe 358
 décroissante 155
 dérivable 150
 dérivable à droite 150
 dérivable à gauche 150
 dérivable sur un intervalle 150
 dérivée 151
 dérivée d'ordre n 151
 développable en série entière 206
 différentiable 238, 242, 246
 différentielle 239
 discontinue 146
 dominée 176
 en escalier 178, 262
 ensemble de définition 174
 équivalente 176
 image 147
 impaire 174
 indéfiniment dérivable 152
 intégrable sur $\mathbb{R}^n$ 267
 intégrable sur un pavé 262
 intégrable sur un segment 179
 limite 146, 235, 236
 localement intégrable 214
 majorée 148
 minorée 148
 monotone 155
 négligeable 176
 numérique 234
 paire 174
 périodique 174
 prolongement par continuité 146
 sinus 158
 sinus hyperbolique 162
 strictement croissante 154
 strictement décroissante 155
 strictement monotone 155
 tangente 159
 tangente hyperbolique 162
 valeur moyenne 180
 vectorielle 235
 de répartition marginales 358

- forme bilinéaire 110
- forme irréductible 35
- forme polaire 110
- forme quadratique 110
 - définie positive 111
 - positive 110
- formule
 - des probabilités totales 303
- formule d'Euler 63
- formule de Grassmann 76
- formule de Hadamard 202
- formule de Moivre 63
- formule de Taylor avec reste intégral 182
- formule de Taylor d'ordre 1 239, 242
- formule du Binôme de Newton 30, 32
- fraction rationnelle 50
- fréquence
 - de la classe $]a_i; a_{i+1}]$ 398
 - de la valeur x_i 398
- fréquence cumulée
 - en a_i 398
 - en x_i 398

G

- gradient 238
- groupe
 - abélien 27
 - commutatif 27

H

- homothétie vectorielle 78
- hyperbole 119
- hyperplan 76

I

- identité du parallélogramme 111
- image 18, 20, 28, 82
- image réciproque 20
- image vectorielle 66
- implication 3
- inclusion 10, 294
- individu 392
- inégalité
 - de Bienaymé-Tchebychev 374
 - de Cauchy-Schwarz 39, 180
 - de la moyenne 180
 - de Markov 374
 - des accroissements finis 243
 - triangulaire 127
 - triangulaire renversée 127

- injectivité 19, 82
- intégrale
 - par rapport à l'abscisse curviligne 270 178
 - absolument convergente 215
 - convergente 214
 - curviligne 270
 - d'une fonction sur un segment 179
 - de surface 274
 - divergente 214
 - sur un pavé de $\mathbb{R}^n$ 262
 - multiples 264
 - triples 264
- intersection 294, 295
- intervalle 127
 - de confiance d'une moyenne 380
 - de probabilité 324
 - fermé borné 127
 - ouvert 127
 - semi-fermés 127
 - semi-ouverts 127
 - stable 143
- inversibilité 95
- isométrie vectorielle 116
- isomorphisme de groupes 27

J

- jacobien 242

L

- ligne de niveau 234
- limite simple 194
- limite uniforme 194
- logique 2
- loi
 - beta 344
 - binomiale 331, 383
 - binomiale négative 335
 - de Bernoulli 330
 - de Dirac 330
 - de Fisher 347
 - de Morgan 3
 - de Poisson 334, 383
 - de Student 346
 - de Weibull 343
 - du Khi-deux 346
 - exponentielle 342
 - géométrique 334
 - gamma 343
 - hypergéométrique 331
 - normale 338

- uniforme continue 342
- uniforme discrète 330
- d'une variable aléatoire réelle 306
- de composition interne
 - associative 26
 - commutative 26
 - distributive 27
 - élément neutre 26
- faible des grands nombres 375
- forte des grands nombres 375
- marginale 351, 354, 359
- longueur 39
- longueur d'une courbe 250

M

- majorant 15, 126
- matrice
 - anti-symétrique 88
 - augmentée 58
 - carrée 86
 - coefficients 86
 - colonne 86
 - d'une application linéaire 90
 - de passage 91
 - des coefficients 58
 - diagonale 86, 88
 - échelonnée 58
 - échelonnée réduite 59
 - équivalente 91
 - inversibilité 88
 - identité 86
 - ligne 86
 - nulle 86
 - produit 87
 - puissance 106
 - somme 87
 - symétrique 88
 - taille 86
 - trace 86
 - transposee 86
 - triangulaire inférieure 88
 - triangulaire supérieure 88
 - de corrélation 364
 - de variance-covariance 363
 - hessienne 246
 - jacobienne 242
 - semblable 92
- maximum global 247
- maximum local 155, 247
- médiane 322
- mesure de probabilité 295

- méthode d'orthogonalisation de Schmidt 115
- méthode de Cramer 98
- mineur 94
- minimum global 247
- minimum local 155, 247
- minorant 15, 126
- modalité 396
- mode 322
- moments d'une variable aléatoire réelle 318
- morphisme d'anneaux 31
- morphisme de groupes 27
- moyenne conditionnelle 421
- moyennes marginales 421

N

- négation 2
- nombre complexe
 - affixe 66
 - conjugué 62
 - forme algébrique 62
 - forme exponentielle 63
 - forme trigonométrique 63
 - image 66
 - imaginaire pur 62
 - module 62
 - partie imaginaire 62
 - partie réelle 62
- nombre premier 23
- nombres rationnels 22
- norme 39
- norme euclidienne 234
- noyau 28, 82

O

- observation 393
- orthogonal 114

P

- parabole 118
- parallélépipède 44
- parallélogramme 44
- partie bornée 15, 126, 234
- partie dense 128
- partie entière 50, 128
- partie majorée 126
- partie minorée 126
- partie principale 50
- partition 11
- partition de l'univers 296
- pas d'une subdivision 178

- permutations 290
- PGCD 34
- pivot de Gauss 58
- plan 39
- plan tangent à une surface 252
- plan vectoriel 76
- plus grand élément 16
- plus petit élément 16
- point
 - anguleux 175
 - asymptotique 260
 - birégulier 251
 - d'inflexion de la courbe, 175 254, 259
 - de rebroussement 176, 255
 - fixe 143
 - fixe attractif 143
 - fixe instable 143
 - fixe répulsif 143
 - fixe stable 143
 - régulier 250, 252, 258
 - selle 247
 - singulier 250, 252
 - stationnaire 247, 259
- pôle 50
- polynôme
 - algébriquement clos 48
 - caractéristique 102
 - coefficient dominant 46
 - coefficients 46
 - degré 46
 - irréductible 47, 48
 - produit 46
 - racine 47
 - racine d'ordre k 47
 - scindé 48
 - somme 46
- population 392
- position pivot 59
- potentiel 271
- PPCM 35
- prédicat 2
- primitive d'une fonction 182
- probabilité 295
 - conditionnelle 302
 - uniforme 298, 300
- produit
 - cartésien 11
 - mixte 40
 - scalaire 38, 111
 - vectoriel 39
- projecteur orthogonal 115
- prolongement 18

Q

- quantificateurs 6
- quantile 323

R

- racine n -ième 63
- raisonnement mathématiques
 - par contraposée 7
 - par contre-exemple 7
 - par déduction 7
 - par l'absurde 7
 - par récurrence 7
- rang 83
- rayon de convergence 202
- rayon de courbure 251
- réciproque 4
- recouvrement 11
- règle de d'Alembert 202
- règle de dérivation des fonctions composées
 - 239
- relation binaire
 - antisymétrique 14
 - ensemble d'arrivée 14
 - ensemble de départ 14
 - graphe 14
 - réflexive 14
 - symétrique 14
 - transitive 14
- relation d'équivalence 15
- relation d'ordre 15
- relation d'ordre total 15
- relation de Chasles 180
- repère 38
- repère orthonormé 42
- repère polaire 258
- reste de Taylor 166
- reste de Young 166
- reste intégral 182
- restriction 18
- rotation 67
- rotationnel 243

S

- segment 127
- série
 - statistique 397
 - absolument convergente 187
 - alternée 188
 - convergente 186
 - de Fourier 210

- de Riemann 186
 - divergente 186
 - double absolument convergente 267
 - double convergente 266
 - exponentielle 186
 - géométrique 186
 - harmonique 186
 - harmonique alternée 186
 - nature 186
 - reste 186, 198
 - semi-convergente 187
 - somme 186, 198
 - somme partielle 186
 - série statistique
 - groupée 398
 - non groupée 398
 - qualitative 399
 - simple 397
 - univariée 397
 - similitude 67
 - sinus 158
 - solutions singulières 223
 - somme de Riemann 180
 - sous-anneau 30
 - sous-corps 32
 - sous-groupe 27
 - sous-suite 135
 - soustraction 294
 - statistique 394
 - descriptive 394
 - inductive 395
 - inférentielle 395
 - statistiques 394
 - subdivision σ de $[a; b]$ 178
 - suite
 - adjacentes 135
 - alternée 131
 - arithmético-géométrique 139
 - bornée 131
 - arithmétique 138
 - bornée 131
 - complexe 130
 - complexe bornée 135
 - complexe convergente 136
 - convergente 134
 - croissante 131
 - décroissante 131
 - divergente 134
 - dominée 136
 - équivalente 136
 - extraite 135
 - géométrique 138
 - majorée 131
 - minorée 131
 - monotone 131
 - négligeable 136
 - numérique 130
 - périodique 131
 - récurrente 130
 - récurrente linéaire 107
 - récurrente linéaire d'ordre deux 142
 - réelle 130
 - stationnaire 132
 - strictement croissante 131
 - strictement décroissante 131
 - strictement monotone 131
 - support d'une subdivision 178
 - surface paramétrée 252
 - surfaces de niveau 234
 - surjectivité 19, 83
 - système complet d'événements 296
 - système de Cramer 99
 - système différentiel linéaire 106
 - système linéaire
 - équivalent 55
 - homogène 54
 - non homogène 54
 - triangulaire 55
- T**
- tableau
 - de données brutes 397
 - de l'analyse de la variance 484
 - statistique 396
 - tableau statistique
 - forme empilée 396
 - tangente à une courbe 254, 258
 - tautologie 4
 - taux de variation d'une fonction 241
 - tétraèdre 44
 - théorème
 - de la limite centrale 378
 - de Bayes 303
 - de Bézout 35
 - de Cayley-Hamilton 102
 - de changement de variables 264
 - de d'Alembert 48
 - de Fubini 264, 267
 - de Gauss 35
 - de Green dans le plan 272
 - de la base incomplète 76
 - de Pythagore 114
 - de Schwartz 247

- de Stokes 275
- du point fixe 244
- du rang 83
- transformation de similitude 103
- transformée de Laplace 52, 278
- translation 67, 78
- travail 271
- triangle 44
- triangle de Pascal 292
- triangularisation 104

U

- union 294, 295
- unité statistique 392
- univers 294, 392

V

- valeur absolue 126
- valeur décimale approchée par défaut 128
- valeur décimale approchée par excès 128
- valeur propre 102

- variable aléatoire réelle 306
 - centrée 319, 324
 - continue 306
 - discrète 306
 - loi 306
 - réduite 318, 319
- variables aléatoires indépendantes 366
- variance 318
 - conditionnelle 421
 - marginale 421
- variance et écart-type 318
- vecteur 38
- vecteur aléatoire 350
- vecteur normal à un plan 43, 252
- vecteur propre 102
- vecteur tangent à une courbe 250, 253
- vecteurs
 - orthogonaux 114
 - orthonormaux 114
- vecteurs colinéaires 39
- vecteurs liés 75
- voisinage 234
- volume d'un pavé compact 262